2010

Evolving Models: A Density-Based Approach to Modeling Sexual Dimorphism and Adaptive Speciation

Audrey Smith
Utah State University

EVOLVING MODELS: A DENSITY-BASED APPROACH TO MODELING SEXUAL DIMORPHISM AND ADAPTIVE SPECIATION

By

Audrey Smith

A report submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Mathematics

UTAH STATE UNIVERSITY
Logan, Utah

2010

Evolving Models: A Density Based Approach To Modeling Sexual Dimorphism and Adaptive Speciation

Audrey Smith*
Department of Mathematics
Utah State University
Logan, UT 84321
audrey.smith@aggiemail.usu.edu

September 15, 2010

Abstract

In this paper, we begin by extending existing deterministic and individual-based ecological models for sexual dimorphism and adaptive speciation into density-based mathematical models describing the density or number of individuals with various trait values, or phenotypes. These density-based models describe the dynamics of a population of males and females using both clonal and sexual reproduction. Each generation, the populations are subject to mating, mutation, and ecological dynamics including intraspecific competition and carrying capacity of the environment. By avoiding individual-based models, we are able to avoid simulations and instead achieve repeatable results.

Implementing these models numerically, we are able to show how clonal reproduction readily results in sexual dimorphism or evolutionary branching. We compare the numerical results to a stability analysis for the case of a monomorphic population and showcase the need for a more complex stability analysis of a non-monomorphic population. We then evaluate the effect of the rate of mutation and, for a population mating sexually, show how the addition of shared material between genders eliminates sexual dimorphism and evolutionary branching whenever the ratio of shared material gets too large.

Adding the restriction that populations at the beginning of each generation follow a Gaussian distribution, we are able to utilize a method called Effective Particle Theory to analytically determine six difference equations relating the population, mean and variance of each gender, to the same parameters in the next generation. Comparing numerical results to the analytic Effective Particle Theory results, we show how the Effective Particle Theory can provide a good approximation of the numeric results in cases where the populations remain unimodal and how the method might be extended to describe bimodal or multimodal populations.

*A Project Report for Utah State University MS Mathematics Degree

List of Figures

Table 1: Table of Symbols

Symbol	Meaning
$c(z, z')$	Competition between an individual with phenotype z and an individual with phenotype z'
f	Ratio of shared genetic material between sexes, ranges from 0 to 1
$K(z)$	Carrying capacity of the environment for an individual with phenotype z
K_0	Max carrying capacity
L	Number of alleles a species has
M_t, N_t	Number of males and females in generation t
$m_t(z), n_t(z)$	Density of males and females with phenotype z at start of generation t
$\tilde{m}_t(z), \tilde{n}_t(z)$	Density of male and female offspring (pre-mutation) with phenotype z in generation t
$\tilde{\tilde{m}}_t(z), \tilde{\tilde{n}}_t(z)$	Density of male and female offspring (post-mutation) with phenotype z in generation t
$n_{eff}(z)$	Effective population an individual with phenotype z experiences
n^*	Equilibrium population
$P(z)$	An individual's probability of survival
r	Average number of offspring per female
x	Phenotype of the mother, ranges from -1 to 1
y	Phenotype of the father, ranges from -1 to 1
z	Phenotype, ranges from -1 to 1
z_0	"Optimal" phenotype
μ	Probability of an allele switching values from -1 to 1 or 1 to -1
μ_m	Mean phenotype of the male population
μ_f	Mean phenotype of the female population
$\Omega(z, z')$	Probability that an individual with phenotype z' would mutate to another phenotype z
σ_c^2	Variance of the competition equation
σ_k^2	Variance of the carrying capacity equation
σ_ω^2	Variance of the probability of mutation
σ_f^2	Variance of the female population
σ_m^2	Variance of the male population

1 Introduction

1.1 Application Background

Over the last 50 years or so, evolutionary ecologists have strongly debated whether adaptive speciation can occur in nature or even in theory. There have been numerous papers written and models created, but scientists have been handicapped when trying to analyze models due to the limitations in what can be done analytically. As a result, many models and results have been based on long and sometimes difficult to reproduce stochastic simulations.

A simpler problem may be to model sexual dimorphism, the systematic physical and behavioral differences between genders, which clearly exists almost everywhere in nature, and attempt to determine the parameters that may prevent or allow that to occur. To do this, we will create a set of density-based mathematical models describing a theoretical population of males and females and their ecological dynamics over a period of generations. These densities will describe the number of individuals with various traits or phenotype values and the models will be adapted from a set of individual-based models created by Bolnick and Doebeli in "Sexual Dimorphism and Adaptive Speciation: Two Sides of the Same Ecological Coin" (Bolnick and Doebeli, 2003). These models will represent a species mating clonally, or asexually, as well as a population using sexual reproduction and will both incorporate mutation as well as ecological effects due to intraspecific competition and carrying capacity.

After developing these models, we will compare their dynamics using numerical methods such as the Fast Fourier Transform and analytical methods such as Effective Particle Theory. Then, using both the numeric and analytic approaches, we will analyze the effect of various parameters and determine differences that may exist between the evaluation methods. Since much work has been done analyzing the effect of intraspecific competition, i.e. competition within the species (Bolnick, 2004; Burger et al., 2006; Slatkin, 1984), we will instead devote some time to analyzing the effect of the mutation parameter and, in the case of a population mating sexually, we will analyze the effect of the ratio of shared genetic material between the

male and female (which determines how offspring phenotype is determined and how much weight each parent has in determining phenotype).

We will also show how the density-based model can be used to predict sexual dimorphism and speciation as in Bolnick and Doebeli (2003). We will perform a stability analysis for a population mating clonally, and finally, by imposing the restriction that the male and female populations follow a Gaussian distribution at the beginning of each generation, we will determine a set of difference equations relating the population, mean and variance of each gender in one generation, to the same parameters in the following generation.

The purpose of creating these density-based models is to show that the ecological dynamics of a species can be determined mathematically using numeric and analytic methods while still allowing a dynamically changing population. A major improvement of the density-based models over individual-based models is the fact that density-based models provide reproducible results and do not require repeated stochastic simulations to determine dynamics. Another advantage of the density-based model is that, in theory, a more general stability analysis could be performed on the difference equations describing the population, mean and variance calculated using Effective Particle Theory. (More general in the sense that the dynamics of a diverse population could be examined versus just a monomorphic, or identical, population.) This stability analysis could in turn provide more insight into the behavior of this complex system.

1.2 Bolnick and Doebeli's model

Slatkin (1984) discussed three main classes of assumptions when it came to explaining sexual dimorphism, (1) differences in the needs of males and females due to energetic needs or gender roles, (2) the existence of more than one optimal trait value for each gender, and (3) intraspecific competition for limited resources. Ignoring the first two classes of assumptions, we will focus on (3), intraspecific competition as the impetus for sexual dimorphism and speciation.

To demonstrate how intraspecific competition could lead to either sexual dimorphism

or adaptive speciation, Bolnick and Doebeli (2003) created a deterministic model and a stochastic model for sexual dimorphism. Each model began with a population of males and females that were characterized by some trait value or phenotype. Then, depending on that trait value and the trait values of the rest of the population, individuals were subject to a dynamically changing fitness landscape determined by the carrying capacity of the environment and the intraspecific competition between individuals. Based on the resources available for a given phenotype and the number of individuals with that same phenotype, each individual in the population was given some probability of surviving to adulthood to reproduce, which then determined the array of phenotypes and the fitness landscape for the next generation of individuals.

Using these models, Bolnick and Doebeli demonstrated how a population generally evolves to the phenotype most suited to use the available resources in an environment. This however, causes intense competition with all the other individuals at that trait value, lowering their probability of survival; whereas at the same time, other trait values may experience less competition and flourish. Theoretically, this phenomenon can lead to speciation or sexual dimorphism depending on mating behavior and the parameter values governing the system.

In order to create models that would mimic the stochastic behavior and dynamic ecological characteristics of the Bolnick and Doebeli (2003) model, we fused the concepts from their deterministic and individual-based models and created a density-based model using clonal reproduction and a second density-based model using sexual reproduction. The model for sexual reproduction is an extension of the clonal model where, rather than clonal reproduction, we assume that each parent will have some effect on the offspring. In these models, rather than describing ecological dynamics for a single individual, we focus on the densities of the male and female populations and describe the dynamics and phenotype distributions of males and females at discrete time increments. Unlike the individual simulation based models, the density-based models can be written down as a set of mathematical equations,

yet they mimic stochastic effects. This means that the results are repeatable and consistent, not requiring multiple lengthy simulations for every parameter combination before any dynamics can be understood, yet the results are more informative than a deterministic model that could only describe dynamics for populations of identical males and females.

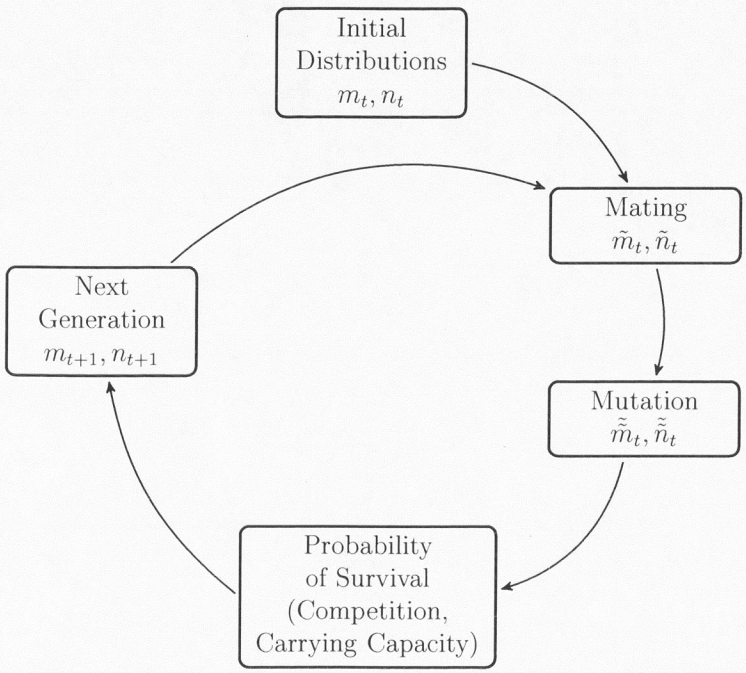

Figure 1: Flowchart of Life Cycle: Each generation, we begin at the top with initial distributions describing the phenotypes of the male and female populations. Then, mating and mutation occur, followed by the effects of competition and carrying capacity. This cycle determines the distribution of phenotypes that survive to the beginning of the next generation, where the cycle begins again.

2 Clonal Model

In this model, we begin generation t with a population of M males and N females who have survived to adulthood but who have not yet mated. As in Bolnick and Doebeli (2003), each individual within the population is characterized by some trait value or phenotype, z, which represents some characteristic like body size. These trait values range from -1 to 1

and determine how the populations interact. From a genetic standpoint, the trait values can be thought of as being determined by adding up the values of each allele at L loci, or gene locations on a chromosome. Each locus is assumed to have equal weight in determining the phenotype and each has two alleles that are assigned a value of either -1 or 1. This sum can then be normalized by dividing by $2L$ to arrive at a phenotype value lying between -1 and 1. Rather than specifically tracking each individual however, we let $m_t(z)$ and $n_t(z)$ be densities describing these populations of males and females. We then focus on these densities as a whole to determine the results of mating, mutation, and the changing fitness landscape.

Each generation begins with a population of males and females that have survived to adulthood but have not mated. From here, clonal reproduction and mutation occur and then, depending on the phenotypes of the population and the carrying capacity of the environment, a certain proportion of these individuals will survive to adulthood giving us new densities $m_{t+1}(z)$ and $n_{t+1}(z)$ representing the males and females at the start of generation $t+1$. These steps are described in more detail below and are shown in Figure 1.

2.1 Clonal Reproduction

Each generation, the populations go through a form of clonal reproduction. We assume that males pass their phenotype to sons and females pass their phenotype to daughters. As in the Bolnick and Doebeli (2003) model, we avoid the possibility of a gender dying out by defining the number of offspring of each gender according to a female-dominant mating rule. This means that the number of offspring of both genders is determined by the number of females who are available to mate and is not affected by the number of males. We let r represent the average number of offspring per mating, and assume that half the offspring will be male and half will be female. This means that the number of new offspring of each gender, $\tilde{M}_t$ and $\tilde{N}_t$, is equal to the number of mating females multiplied by the average number of offspring r divided by two. That is,

$$\tilde{M}_t = \tilde{N}_t = \frac{rN_t}{2}, \tag{1}$$

and the corresponding densities of male and female offspring following mating are denoted by $\tilde{m}_t(z)$ and $\tilde{n}_t(z)$.

2.2 Mutation

From a theoretical standpoint, the genotype of the offspring in this clonal reproduction model should be the same as the genotype of the parent. In nature however, errors or mutation may occur during the reproductive process, leading to small changes in the genotype and possibly the phenotype of the offspring.

To represent mutation mathematically, it was necessary to allow slight changes in the phenotype or trait value of some offspring to represent the errors that can occur during the reproductive process. This leads to a small number of male and female offspring with slightly different trait values than their parents. While theoretically occurring during or before the mating process, it was simpler mathematically to implement mating and mutation separately.

Bolnick and Doebeli (2003), in their individual-based model, implemented mutation at the genotype level. Individuals were assumed to be diploid (having two sets of chromosomes) with L loci and each locus had two alleles, each assigned a value of 1 or -1 where each locus was also assumed to have an equal effect on the phenotype. To create small changes to the phenotype, they used a Bernoulli trial to give each allele the probability $\mu = .001$ of reversing its value due to mutation. This was being applied to $2L$ alleles, so together a binomial distribution could describe the number of alleles that had switched values. After applying this probability of reversing value to each allele, the values were summed up and normalized to determine the post mutation offspring phenotype.

In our model, to represent mutation's effect on the phenotype, we assumed that the number of alleles, $2L$, was large enough that we could represent the effect of mutation using a continuous distribution. This distribution would represent the probability of an allele switching values from -1 to 1 or from 1 to -1; an effective change to the unnormalized sum of +2 or -2 for each mutated allele. Since either occurrence was assumed to be equally likely, the mean effect of mutation was zero. We therefore chose to represent mutation using a

Gaussian distribution

$$\Omega(z, z') = \frac{1}{\sqrt{2\pi\sigma_\omega^2}} \exp\left[-\frac{(z - z')^2}{2\sigma_\omega^2}\right], \tag{2}$$

where $\Omega(z, z')$ represented the probability that an individual with phenotype z' would mutate to some other phenotype z.

In (2), the analogue to the probability of mutation, μ, in the individual-based model is σ_ω. This value determines how wide the mutation distribution is, which determines how likely it is to alter an offspring phenotype. Higher values of σ_ω indicate more mutation (i.e. large μ) whereas, as $\sigma_\omega \to 0$, the rate, or probability of mutation also goes to zero ($\mu \to 0$). The extent to which a particular phenotype is altered by mutation, given a particular σ_ω, is determined by the phenotypic distance between z' and z; an individual will be much more likely to mutate to a nearby phenotype than a distant one.

To demonstrate how this Gaussian mutation distribution can be used to cause slight changes to the phenotypes of a small number of males and females, we can look at things from a probabilistic standpoint. Using the definition of our mutation distribution, we know that

$$\Omega(z, z') = P(\text{mutate from } z' \text{ to } z),$$

and

$$\tilde{m}_t(z')dz' = \#\text{ of individuals with phenotype } z',$$

so

$$P(\text{mutate from } z' \text{ to } z) \cdot (\#\text{ of individuals with phenotype } z') = \Omega(z, z') \cdot \tilde{m}_t(z')dz',$$

would equal the expected number of individuals mutating from z' to z. To find the total number of individuals who have some phenotype z after mutation, we sum over the entire range of z' using an integral,

$$\tilde{\tilde{m}}_t(z) = \int \Omega(z, z') \cdot \tilde{m}_t(z')dz', \tag{3}$$

which represents the density of male offspring following mutation in generation t. Similarly,

$$\tilde{\tilde{n}}_t(z) = \int \Omega(z, z') \cdot \tilde{n}_t(z')dz', \tag{4}$$

describes the female density after mutation.

The effect of mutation on a population can be visualized in Figure 2 in which a Gaussian male density is plotted prior to and following convolution with an exaggerated mutation distribution. This image shows that the Gaussian mutation distribution had the effect of spreading the male distribution slightly, lowering its peak and widening its tails indicating that there was some spread in the phenotypes, or increase in variance, of the males.

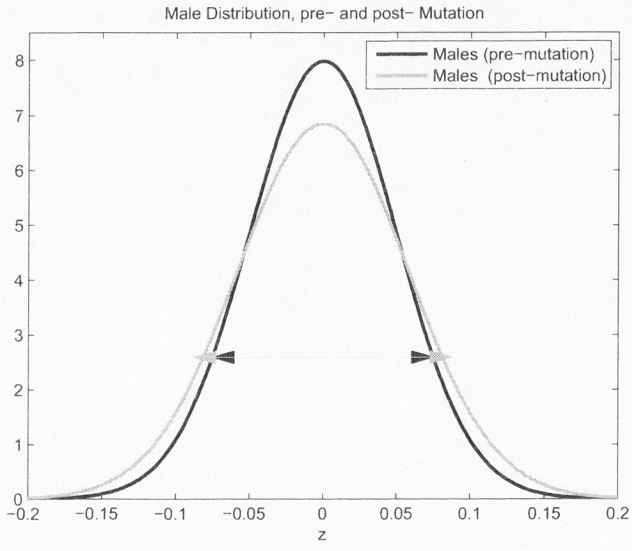

Figure 2: Male Distribution pre- and post- mutation. Note that mutation has the effect of lowering the peak of the distribution and widening the tails, in effect, spreading out the range of phenotypes.

2.2.1 Numerical Implementation of Mutation using FFT

Equations (3) and (4) describe the densities of males and females following mating and mutation. In order to determine what these densities are however, we needed to implement mutation numerically as a convolution in MATLAB using a Fast Fourier Transform (FFT). In MATLAB, this operation is done using vectors to describe the distributions $\tilde{m}_t$, $\tilde{n}_t$ and Ω, and then transforming these vectors into Fourier space using the FFT function. The Fourier

Transform of a function is defined as

$$\hat{f}(k) = \int_{\infty}^{\infty} f(z)e^{-ikz}dz,$$

and a nice property is that the Fourier Transform of the convolution is proportional to the product of the Fourier Transform of the components in the convolution (Logan, 2006), so

$$\widehat{(\tilde{m}_t * \Omega)}(k) = \frac{1}{2\pi}\hat{m}_t(k)\hat{\Omega}(k).$$

Using this property, we obtain the value of the convolution by computing the product of the components in Fourier space, then taking the inverse of the Fourier Transform. This gives the male and female densities following mutation, $\tilde{\tilde{m}}_t(z)$ and $\tilde{\tilde{n}}_t(z)$.

2.3 Probability of Survival

After determining these post-mutation densities, we needed to determine the fitness landscape which would affect which offspring survived to adulthood to produce the next generation of offspring. The fitness landscape describes how hospitable the environment is to each phenotype and is affected both by the ecological environment as well as the density of each phenotype in the population. This is represented as a probability of survival for each phenotype. Here we used a logistic probability of survival equation:

$$P(z) = 1 - \frac{r-1}{r}\frac{n_{eff}(z)}{K(z)}, \tag{5}$$

which was a departure from Bolnick and Doebeli (2003), who used a modification of the Beverton-Holt equation,

$$P(z) = \frac{1}{1 + \dfrac{r-1}{K(z)}n_{eff}(z)}. \tag{6}$$

This was a necessary change because we wished to, with certain restrictions, be able to analytically compute the probability of an offspring surviving to adulthood. This required computation of the convolution of $P(z)$ with $\tilde{\tilde{m}}_t(z)$ and $P(z)$ with $\tilde{\tilde{n}}_t(z)$, which, due to the form of (6) would result in an intractable integral. By using equation (5), we are able to compute these integrals using the properties of Gaussian distributions and while these

equations appear quite different, they do have several similar properties. Both equations give the probability that a young offspring with phenotype z will survive to adulthood and produce offspring themselves. In each equation, r represents the average number of offspring per female, $n_{eff}(z)$ represents the effective population size that an individual with trait value z experiences, and $K(z)$ represents the carrying capacity, or the maximum population that the environment can continually sustain. Equations (5) and (6) both also have the property that for small populations (when the population is much smaller than the carrying capacity), they have the same linear growth rate, r. Neglecting z-dependence and allowing n to represent the population, these equations both yield the fixed point

$$n_{t+1} = n_t = n^* = K,$$

which is stable when r is between 1 and 2. Additionally, if r is close to 1, then (5) is similar to the first two terms of a Taylor expansion of (6). A major difference between the equations is the fact that for large r, the logistic equation can be destabilized whereas the Beverton-Holt equation is always stable. For this reason, we restricted r to be between 2 and 4 for all simulations.

2.3.1 Effective Population and Competition

In order to utilize (5), we must determine the components within it. The effective population with which an individual must compete, $n_{eff}(z)$, is based on how much it must compete with other members of the population for limiting resources such as food or breeding sites. This is determined using the symmetric competition equation

$$c(z, z') = \exp\left\{-\frac{(z - z')^2}{2\sigma_c^2}\right\}, \tag{7}$$

as in the Bolnick and Doebeli (2003) model. In this equation, $c(z, z')$ describes the strength of competition between an individual with phenotype z and another individual with phenotype z'. A graph representing the competition of an individual with phenotype $z = 0$ against individuals with phenotypes z' is shown in Figure 3. This figure shows how the strength of

competition is affected both by the phenotypic distance between individuals and by σ_c, which is inversely related to the strength of competition. The wider the competition distribution is, the wider the phenotype range an individual must compete heavily with and, assuming the curve stays the same, the further their offspring will have to evolve to escape its effect.

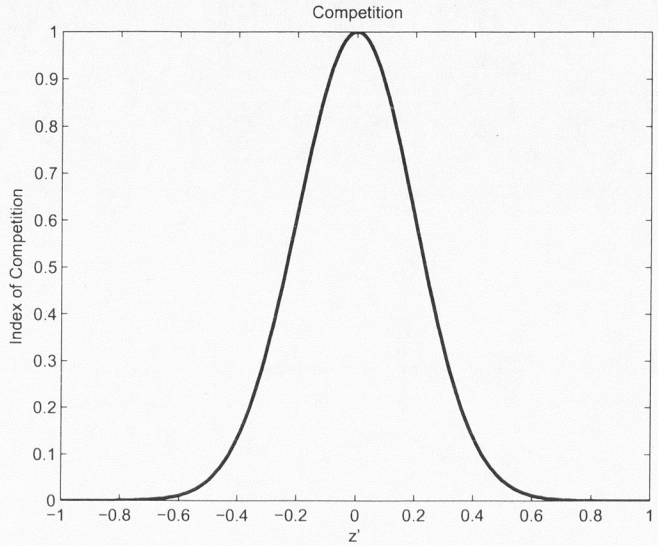

Figure 3: Competition Curve of individual with phenotype $z = 0$ when $\sigma_c = 0.2$. Note how competition is most intense with individuals of the same or similar phenotype.

From (7), we see that individuals with the same phenotype will experience a competition value of 1 meaning that they will have the largest possible effect on each other. This is because they are suited for the same food sources or habitat. On the other hand, individuals with significantly different phenotypes may be suited for different food sources and consequently will experience less competition with each other. An example of this may be an herbivore such as a giraffe, where the trait value z represents height. Individuals with the same trait value, or height, will compete directly with each other because they are both eating food that grows at the same level, whereas another individual who is substantially taller or shorter, may not compete as much because they may primarily eat food that grows at a different height.

Using this competition equation and the notion that competition between different members of a population will vary according to trait value, the effective population an individual with phenotype z experiences can be computed by calculating the competition coefficient for some phenotype z' and then calculating the number of individuals with that phenotype. As in Bolnick and Doebeli (2003), we assume that there is no difference between genders and that the effect of a male with phenotype z' will be the same as the effect of a female with phenotype z'. So, to compute the total effective population the individual experiences, we sum over the entire range of phenotypes, i.e.

$$n_{eff}(z) = \int c(z, z') \left[\tilde{\tilde{n}}_t(z') + \tilde{\tilde{m}}_t(z')\right] dz', \tag{8}$$

where $\tilde{\tilde{m}}_t(z')$ and $\tilde{\tilde{n}}_t(z')$ represent the density of males and females, respectively, with phenotype z'.

So, if all individuals in the population (male and female) are identical, then their competition coefficient will be 1. This means that the effective competition each individual will experience will be the entire population of males and females. If, on the other hand, there is some variance in the population, then the effective population will be less.

2.3.2 Carrying Capacity

The effective population and competition explain how the members of a population affect each other, the other main factor affecting an individual's probability of survival however, is the environment. Specifically, how many individuals of a particular phenotype can the environment sustain? To model this, we use an equation for carrying capacity

$$K(z) = K_0 \exp\left\{-\frac{(z - z_0)^2}{2\sigma_k^2}\right\}, \tag{9}$$

where K_0 represents the maximum carrying capacity and z_0 is some optimal phenotype value where resources are most abundant. This value was arbitrarily set to equal zero as in Bolnick and Doebeli (2003) and, using the giraffe example, this optimal trait could represent the height that would allow an individual to access the most food. The last parameter, σ_k,

determines how fast the abundance of resources decreases as we move away from the optimal trait value. We note here that while many evolutionary biologists use the assumption that there is a separate ideal trait for males and females (Slatkin, 1984), we have chosen to simplify the mathematics and focus on the effects of competition by assuming that there is a single ideal trait for both genders.

As with competition, the value of $K(z)$ is affected by the phenotypic distance between z and z_0. If $z = z_0$, the "optimal" trait, then $K(z) = K_0$, whereas with increasing distance from the optimal trait, then $K(z)$ decreases from K_0 meaning that the individuals are not as suited for the environment and fewer can be sustained.

Putting this together, we see that the higher the ratio $\frac{n_{eff}(z)}{K(z)}$, the larger the second term in (5) and the lower the probability of survival. To visualize the effect of the probability of survival equation, we plotted the probability of survival curve with the corresponding scaled male and female densities for three different scenarios in Figure 4. These figures show how the probability of survival is affected by the density of individuals: low population densities imply higher $P(z)$ and high population densities imply lower $P(z)$. We also see that this probability is affected by the phenotype value itself, higher near the optimal trait z_0 and lower near the extreme phenotypes (1 and -1), though the phenotype appears to have less impact than density.

To implement the effects of (5) numerically, we computed (8) as a convolution using another FFT in MATLAB. Finally, to represent the populations after being acted upon by probability of survival, we multiply the male and female post-mutation densities by the probability of survival equation to arrive at the next generation of males and females who have survived to adulthood and are ready to mate,

$$m_{t+1}(z) = P(z)\tilde{\tilde{m}}_t(z) \tag{10}$$

and

$$n_{t+1}(z) = P(z)\tilde{\tilde{n}}_t(z), \tag{11}$$

and the process repeats.

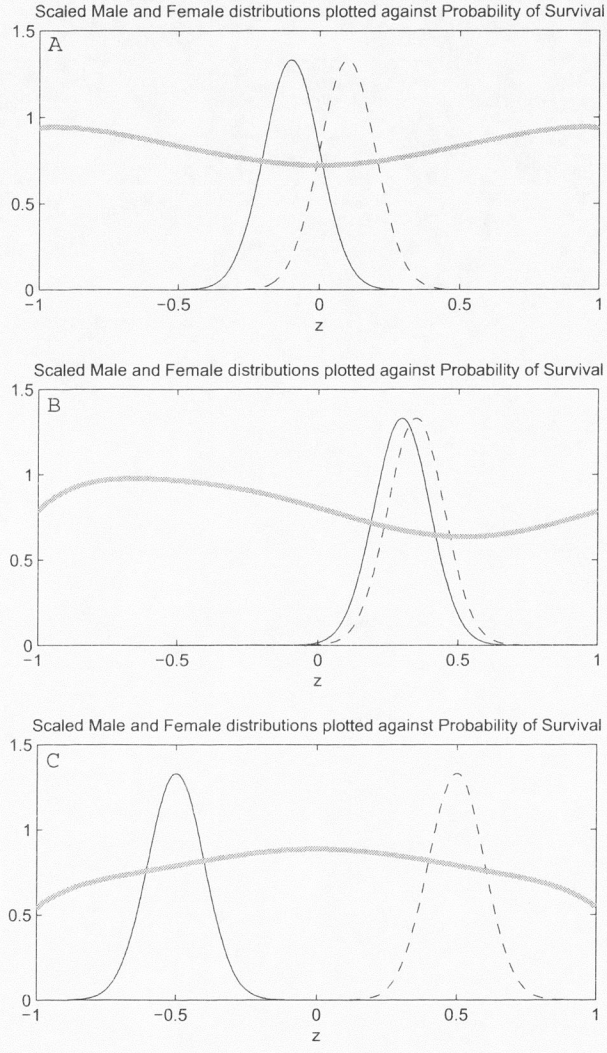

Figure 4: Probability of Survival curve (gray) and corresponding male (black solid) and female distributions (black dashed) for A) $\mu_m = -0.5$, $\mu_f = 0.5$, B) $\mu_m = 0.3$, $\mu_f = 0.35$, and C) $\mu_m = -0.1$, $\mu_f = 0.1$ when $\sigma_m = \sigma_f = 0.15$, $K = 2000$, $\sigma_c = 0.5$ and $\sigma_k = .6$. Note that areas of higher population density correspond to lower probabilities of survival, as do the edges of the phenotype range.

3 Stability Analysis

Here we provide a stability analysis of our clonal model to determine situations leading to dimorphism and speciation. We will assume a monomorphic population and look at the

dynamics of rare mutants in the population. For this analysis, we begin with the assumption that males and females are each monomorphic for some trait value z_m and z_f respectively, so that all males are identical and all females are identical and their distributions can be described by delta functions,

$$m_t(z) = M_t \delta(z - z_m)$$

and

$$n_t(z) = N_t \delta(z - z_f),$$

where we have used a slight departure of notation. In this stability analysis: M_t and N_t represent the populations of males and females after mating but before competition.

To determine the number of male offspring that survive to adulthood, we multiply the number of male offspring by the probability of survival computed from (5). Because males are monomorphic for some trait z_m, then the competition coefficient between the males, $c(z_m, z_m)$, is one. This means that the effective population of males each male experiences is equal to the total number of males, M_t. The effective population of females each male experiences is computed by multiplying the competition coefficient, $c(z_m, z_f)$, representing the strength of competition between males and females, by the total number of females, N_t. This means that the population of males after death due to competition but before reproduction is

$$\tilde{M}_t = M_t P(z) = M_t \left(1 - \frac{r-1}{r} \frac{M_t + c(z_m, z_f) N_t}{K(z_m)}\right),$$

where r again represents the average number of offspring per mating. Similarly, the number of female offspring who survive to adulthood to mate is given by

$$\tilde{N}_t = N_t P(z) = N_t \left(1 - \frac{r-1}{r} \frac{N_t + c(z_f, z_m) M_t}{K(z_f)}\right).$$

Now, assuming clonal mating, we can determine the number of offspring at the start of generation $t+1$ using the female dominant mating rule described in the Section 2. The number of new males and females is

$$M_{t+1} = N_{t+1} = \frac{r}{2} \tilde{N}_t,$$

which matches (1). Then, using this result, a difference equation relating the number of males and females in generation t to the number of males and females in generation $t+1$ is obtained. Males and females remain monomorphic with trait values z_m and z_f respectively,

$$M_{t+1} = N_{t+1} = \frac{r}{2} N_t \left(1 - \frac{r-1}{r} \frac{N_t + c(z_f, z_m) M_t}{K(z_f)}\right). \tag{12}$$

Solving this population equation for a fixed point, we obtain

$$N^* = M^* = \frac{r-2}{r-1} \frac{K(z_f)}{1 + c(z_m, z_f)}, \tag{13}$$

which represents the equilibrium population numbers for males and females. In this equation, because we assume an equilibrium population, we must restrict r so that it is greater than or equal to 2 to ensure that the population is sustaining itself.

After determining these equilibrium points, we will now consider the fitness and explore the population dynamics of rare mutants attempting to invade this steady population. Assuming the overall population has already reached its equilibrium population, N^*, then the density of female mutants with phenotype $z_{f'}$ in generation $t+1$ is given by the equation

$$N'_{t+1} = \frac{r}{2} N'_t \left(1 - \frac{r-1}{r} \frac{N^* \left(c(z_{f'}, z_f) + c(z_{f'}, z_m)\right)}{K(z_{f'})}\right),$$

meaning that the growth rate of rare female mutants is

$$w_f(z_{f'}, z_f, z_m) = \frac{r}{2} \left(1 - \frac{r-1}{r} \frac{N^* \left(c(z_{f'}, z_f) + c(z_{f'}, z_m)\right)}{K(z_{f'})}\right). \tag{14}$$

Similarly, the growth rate of rare male mutants is

$$w_m(z_{m'}, z_f, z_m) = \frac{r}{2} \left(1 - \frac{r-1}{r} \frac{N^* \left(c(z_{m'}, z_f) + c(z_{m'}, z_m)\right)}{K(z_{m'})}\right). \tag{15}$$

In the above equations, $w_f(z_f^*, z_f, z_m)$ and $w_m(z_m^*, z_f, z_m)$ describe the fitness of an individual with phenotype z_f^* or z_m^* in the population. Using these fitness equations, we can determine what happens to the average mean of the population from one generation to the next. Suppose at the start of generation t there are N females with phenotype z_f and n

rare mutant females with phenotype $z_{f'}$, where $N >> n$. That means that the mean of the female phenotypes in generation t is given by

$$\bar{z}_{f,t} = \frac{z_f N + z_{f'} n}{N + n}.$$

and the mean female phenotype in the next generation can be described by

$$\bar{z}_{f,t+1} = \frac{z_f\, N\, w_f(z_f, z_f, z_m) + z_{f'}\, n\, w_f(z_{f'}, z_f, z_m)}{N\, w_f(z_f, z_f, z_m) + n\, w_f(z_{f'}, z_f, z_m)}.$$

where we have used the female fitness function to determine the number of females with phenotype z_f and the number of mutant females respectively. Using the fact that the females with mean z_f are at equilibrium, then we know that $w_f(z_f, z_m, z_f) = 1$ and the expression for the mean in generation $t + 1$ can be simplified to

$$\bar{z}_{f,t+1} = \frac{z_f\, N + z_{f'}\, n\, w_f(z_{f'}, z_f, z_m)}{N + n\, w_f(z_{f'}, z_f, z_m)}.$$

Now, using a Taylor expansion on the female fitness function, we see that

$$w_f(z_{f'}, z_f, z_m) \doteq w_f(z_f, z_f, z_m) + (z_{f'} - z_f) \left.\frac{\partial w_f}{\partial z_{f'}}\right|^{z_{f'}=z_f} = 1 + (z_{f'} - z_f) \left.\frac{\partial w_f}{\partial z_{f'}}\right|^{z_{f'}=z_f},$$

again using the fact that at equilibrium, the fitness function is 1. Analysing the right hand side of this equation, we see that the partial derivative term is actually a selection gradient for the female trait in a stable population with rare mutants. The selection gradient for the male trait can be found in a similar manner by taking the partial derivative of (15) with respect to the mutant phenotype, $z_{m'}$, and then replacing the mutant phenotype value with the trait values of the stable population, z_m. That is,

$$g_f(z_m, z_f) = \left.\frac{\partial w_f}{\partial z_{f'}}\right|^{z_{f'}=z_f}$$

and

$$g_m(z_m, z_f) = \left.\frac{\partial w_m}{\partial z_{m'}}\right|^{z_{m'}=z_m}.$$

Evaluating the partial derivatives, we obtain

$$g_f(z_m, z_f) = -\frac{(r-2)e^{\frac{-z_f^2+z_m^2}{2\sigma_k^2}}\left(z_f\sigma_k^2 + z_m\left(\sigma_c^2\left(1 + e^{\frac{(z_f-z_m)^2}{2\sigma_c^2}}\right) - \sigma_k^2\right)\right)}{2\sigma_c^2\sigma_k^2\left(1 + e^{\frac{(z_f-z_m)^2}{2\sigma_c^2}}\right)} \tag{16}$$

and

$$g_m(z_m, z_f) = \frac{(r-2)}{2}\left(\frac{z_f - z_m}{\sigma_c^2\left(1 + e^{\frac{(z_f - z_m)^2}{2\sigma_c^2}}\right)} - \frac{z_f}{\sigma_k^2}\right). \tag{17}$$

From here, we can use theory from quantitative genetics to hypothesize that over time, the trait value of the males and females will move toward regions of higher fitness determined by the gradients (16) and (17). Referring back to Section 2, we expect these gradients to move the populations to phenotypes ranges with the best combination of low competition and high carrying capacity so that the populations and their offspring will have the highest possible probability of survival.

To actually evaluate how the population means change over time, we can use the geometric series expansion on the expressions for the mean female phenotype in generation t and generation $t+1$, simplify and then determine an expression for the change in the mean. These expansions for the means are shown below.

$$\begin{aligned}\bar{z}_{f,t} = \frac{z_f + z_{f'}\frac{n}{N}}{1 + \frac{n}{N}} &\doteq \left(z_f + z_{f'}\frac{n}{N}\right)\left(1 - \frac{n}{N}\right) \\ &\doteq z_f + (z_{f'} - z_f)\frac{n}{N},\end{aligned}$$

and

$$\begin{aligned}\bar{z}_{f,t+1} = \frac{z_f + z_{f'}\frac{n}{N}\left[1 + (z_{f'} - z_f)g_f\right]}{1 + \frac{n}{N}\left[1 + (z_{f'} - z_f)g_f\right]} &\doteq \left(z_f + z_{f'}\frac{n}{N}\left[1 + (z_{f'} - z_f)g_f\right]\right)\left(1 - \frac{n}{N}\left[1 + (z_{f'} - z_f)g_f\right]\right) \\ &\doteq z_f + (z_{f'} - z_f)\left[1 + (z_{f'} - z_f)g_f\right]\frac{n}{N}.\end{aligned}$$

Now, using these expansions, we determine that the change in the mean over time is given by,

$$\begin{aligned}\frac{\Delta\bar{z}}{\Delta t} &= \bar{z}_{f,t} - \bar{z}_{f,t+1} \\ &= \frac{(z_{f'} - z_f)^2}{\Delta t}\frac{n}{N}g_f,\end{aligned}$$

and, if we make the assumption that the difference between the mutant phenotype and the mean phenotype is proportional to the square root of the timestep, that is $z_{f'} - z_f = \beta\sqrt{\Delta t}$,

then we can define $M_f = \beta^2 \frac{n}{N}$. This assumption is not unrealistic biologically, for example, if we model the spread of phenotypes due to mutation using a random walk diffusion model, then it can be shown that the rate of change of the phenotypes is proportional the square root of the product of the diffusion coefficient, μ, and the timestep, Δt, (that is $z_{f'} - z_f \propto \sqrt{\mu \Delta t}$). So using this relationship, the change in trait values for the females can be described by

$$\frac{\Delta z_f}{\Delta t} = M_f \, g_f(z_m, z_f), \tag{18}$$

and using a similar substitution, the change in trait values for the males can be described by

$$\frac{\Delta z_m}{\Delta t} = M_m \, g_m(z_m, z_f). \tag{19}$$

So we see, the rate of change of the phenotype value is affected by the selection gradients g_f and g_m as well as by the positive constants M_f and M_m which factor in heritability, the rate of mutation, and the length of generations. Proceeding with the stability analysis, we wish to determine the stability of the equilibrium phenotype values. The stability of these phenotype values can be determined by by the sign of the eigenvalues of the matrix on the right hand side of

$$\begin{pmatrix} \dot{z}_m \\ \dot{z}_f \end{pmatrix} = \begin{pmatrix} M_m \dfrac{\partial g_m}{\partial z_m}(z_m^*, z_f^*) & M_m \dfrac{\partial g_m}{\partial z_f}(z_m^*, z_f^*) \\ M_f \dfrac{\partial g_f}{\partial z_m}(z_m^*, z_f^*) & M_f \dfrac{\partial g_f}{\partial z_f}(z_m^*, z_f^*) \end{pmatrix} \begin{pmatrix} z_m \\ z_f \end{pmatrix}.$$

Using properties of linear algebra, we see that the right hand side can be rewritten as product of a diagonal matrix, the Jacobian matrix and the phenotype vector,

$$\begin{pmatrix} \dot{z}_m \\ \dot{z}_f \end{pmatrix} = \begin{pmatrix} M_m & 0 \\ 0 & M_f \end{pmatrix} \begin{pmatrix} \dfrac{\partial g_m}{\partial z_m}(z_m^*, z_f^*) & \dfrac{\partial g_m}{\partial z_f}(z_m^*, z_f^*) \\ \dfrac{\partial g_f}{\partial z_m}(z_m^*, z_f^*) & \dfrac{\partial g_f}{\partial z_f}(z_m^*, z_f^*) \end{pmatrix} \begin{pmatrix} z_m \\ z_f \end{pmatrix} = \begin{pmatrix} M_m & 0 \\ 0 & M_f \end{pmatrix} \mathbf{J} \begin{pmatrix} z_m \\ z_f \end{pmatrix}.$$

Since M_m and M_f are positive quantities, their presence will not affect the sign of the eigenvalues, only their size. For this reason, we can either allow $M_m = M_f = 1$, as in Bolnick and Doebeli (2003), or simply compute the eigenvalues of the Jacobian matrix to simplify the stability analysis.

Equations (18) and (19) describe the rate of change of the male and female phenotypes over time, so we know that the phenotypes will remain stable when these rates are zero. As in Bolnick and Doebeli (2003), it is clear from (16) and (17) that $(z_m^*, z_f^*) = (0, 0)$ will be an equilibrium point resulting in zero rate of change for both phenotypes. Furthermore, assuming that the fixed points are antisymmetric, i.e. $z_m^* = -z_f^*$, we find that

$$z_m^* = -z_f^* = \pm\sigma_c\sqrt{\frac{1}{2}\log\left(-1 + \frac{2\sigma_k^2}{\sigma_c^2}\right)}, \tag{20}$$

is another equilibrium point which exists only when the quantity underneath the radical is positive, that is, when $\sigma_c < \sigma_k$. This second set of fixed points represents dimorphism.

Computing the Jacobian matrix of equations (18) and (19) evaluated at the equilibrium point (z_m^*, z_f^*),

$$\mathbf{J}(z_m^*, z_f^*) = \begin{pmatrix} \dfrac{\partial g_m}{\partial z_m}(z_m^*, z_f^*) & \dfrac{\partial g_m}{\partial z_f}(z_m^*, z_f^*) \\ \dfrac{\partial g_f}{\partial z_m}(z_m^*, z_f^*) & \dfrac{\partial g_f}{\partial z_f}(z_m^*, z_f^*) \end{pmatrix}.$$

we can now find the eigenvalues λ_1 and λ_2 and evaluate their stability. For the fixed point $(0, 0)$, we find that

$$\lambda_1 = -\frac{r-2}{2\sigma_k^2} \quad \text{and} \quad \lambda_2 = -\frac{(r-2)(\sigma_c^2 - \sigma_k^2)}{2\sigma_c^2\sigma_k^2}.$$

Thus $(0, 0)$ is stable when $r > 2$ and $\sigma_c > \sigma_k$. Conversely, this point is unstable when $\sigma_c < \sigma_k$. For the second fixed point (20), we find that the eigenvalues are

$$\lambda_1 = -\frac{r-2}{2\sigma_k^2} \quad \text{and} \quad \lambda_2 = \frac{(r-2)\left(\sigma_c^2 - 2\sigma_k^2\right)\log\left(-1 + \frac{2\sigma_k^2}{\sigma_c^2}\right)}{2\sigma_k^4},$$

indicating that this fixed point is also stable as long as $r > 2$ and $\sigma_c^2 < 2\sigma_k^2$.

Biologically, this means that if the range of phenotypes the environment can sustain is limited, for example if food is only available at a certain height, then over time, there will be selective pressure for the population means to tend toward that height so that more individuals will be able to obtain food. In this case, we assume that carrying capacity has a bigger effect than competition, $\sigma_c > \sigma_k$, causing both males and females to tend toward the phenotype best suited for the environment, $(0, 0)$. If, on the other hand, the environment

can sustain a wide range of phenotypes and competition between individuals of the same or similar phenotype is strong, $\sigma_c < \sigma_k$, then there will be selective pressure for the population to tend toward the antisymmetric fixed points causing the males and females to be different from each other. This will reduce the density of individuals in any one phenotype range allowing them to avoid excessive competition.

These fixed points help predict parameter combinations leading to sexual dimorphism and speciation for a population mating clonally. To analyze what would happen to a population with shared genes between the males and females, we need to extend the model to allow for sexual reproduction.

4 Mating - Sexual Dimorphism

The second model we developed is an expansion of the clonal model that incorporates sexual reproduction and the possibility for sexual dimorphism. To model this, we needed to modify our assumptions slightly. We again used the assumption that each individual is characterized by some trait value ranging from -1 to 1 and that in generation t, there are M_t males and N_t females who have survived to adulthood and are available to mate.

To create a mechanism for realistic mating, we also defined a parameter f, as in the Bolnick and Doebeli (2003) model, to represent the ratio of shared material between the genders. From a genotype approach, f is defined to be the number of shared alleles, N_{shared}, divided by the total number of alleles that a species has, N_{total}, that is,

$$f = \frac{N_{\text{shared}}}{N_{\text{total}}},$$

where we are assuming that males and females have the same number of alleles, so that N_{total} is the same regardless of gender and f will have a constant value.

4.1 Offspring Phenotypes

As in the clonal model, we begin generation t with a population of males and females who have survived until adulthood but who have not yet mated. These populations are again

described by densities $m_t(z)$ and $n_t(z)$, and each generation, the populations reproduce and mutate, then survive to adulthood with some probability based on the fitness landscape. In this model, the populations reproduce sexually meaning that the phenotype of the offspring may not be identical to the phenotype of the parents. The phenotype of the offspring can be determined using a weighted average of the parent phenotypes,

$$z_m = w_1 y + w_2 x \quad \text{and} \quad z_f = w_1 x + w_2 y,$$

where z_m and z_f in the above equations represent the offspring phenotype, x and y represent the phenotype of the mother and father respectively, and w_1 and w_2 are weights that must sum to 1. Because we assumed that a female offspring would be more influenced by the phenotype of the mother and similarly, a male offspring would be more influenced by the phenotype of the father, we decided to utilize the parameter f representing the ratio of shared material to determine our weights. The parent of the same gender can contribute all of their genes to the offspring while the parent of the opposite gender can contribute only the proportion f of their genes. Normalizing these quantities so that the offspring phenotype is not overdetermined, we obtain the weights $w_1 = \frac{1}{1+f}$ and $w_2 = \frac{f}{1+f}$. Thus

$$z_m = \frac{1}{1+f} y + \frac{f}{1+f} x = \frac{y + fx}{1+f}, \tag{21}$$

and

$$z_f = \frac{1}{1+f} x + \frac{f}{1+f} y = \frac{x + fy}{1+f}, \tag{22}$$

describe the phenotype of the offspring given the parent phenotypes x and y and the ratio of shared material between genders, f.

When $f = 0$, the phenotype of the offspring will be solely determined by the parent of the same gender (as in clonal mating). On the other hand, if all material is shared between the sexes, then $f = 1$, and we see that the offspring phenotypes are simply the average of both parent's phenotypes

$$z_m = \frac{y + 1 \cdot x}{1+1} = \frac{y+x}{2}$$

and

$$z_f = \frac{x + 1 \cdot y}{1 + 1} = \frac{x + y}{2},$$

where both parents have equal weight in determining the phenotype of the offspring.

4.2 Mating/Reproduction

To determine how mating influences the evolution of the distributions we must work backwards and determine the parent combinations that could lead to various offspring phenotypes. Solving (21) for y, we see that in order to have a male offspring with phenotype z_m and a mother with phenotype x, then the father needed to have a phenotype value of

$$y = -fx + (1 + f)z_m. \tag{23}$$

Similarly to obtain a female offspring with phenotype z_f where the phenotype of the father is y, then the phenotype of the mother needed to be

$$x = -fy + (1 + f)z_f, \tag{24}$$

which was obtained by solving (22) for x.

Fixing z_m, z_f and f, we see that these are just linear equations with slope equal to $-f$ and intercepts at $1 + f$ times the offspring phenotype. The slope and intercept of this line have the effect of limiting or extending the range of parent phenotype combinations that can produce a specific offspring. A graphical representation of this is shown in Figure 5 which shows the possible combinations of parent phenotypes that produce a male offspring with phenotype $z_m = 0.37$ given that $f = 0.3$.

Now, assuming the range of offspring phenotypes is the same as the range of the parent's, from -1 to 1, we need to determine the number of ways that a particular offspring phenotype could occur along that range. For this, we can look at probabilities. The probability of producing a male offspring with phenotype z_m is proportional to the conditional probability of a female with phenotype x encountering a male with phenotype y as in equation (24). Assuming these events are independent, we get

$$P(z_m) \propto P(x)\, P(y),$$

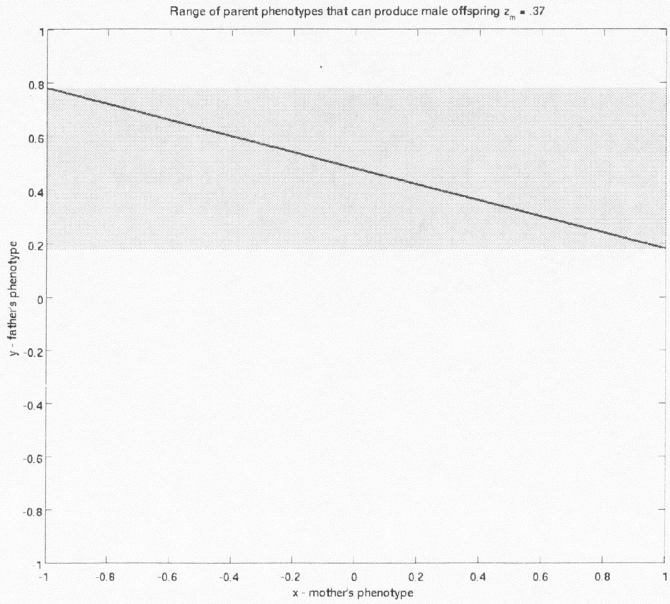

Figure 5: The possible parent phenotype combinations that can lead to male offspring with phenotype $z_m = 0.37$ given $f = 0.3$ lie along the line $y = 1.3z_m - .3x$. Note that the range of male phenotypes that can produce this offspring is limited. Increasing the ratio of shared material, f, would have the effect of increasing the slope of the line and extending the range of possible fathers.

where

$$P(x) = \frac{n(x)}{N}\,dx,$$

and, using a change of variables to convert from y to z_m as in equation (23),

$$P(y) = \frac{m(y)}{M}\,dy = \frac{1}{M}\,m(-fx + (1+f)z_m)\,(1+f)\,dz_m,$$

where the $(1+f)$ term is a result of changing variables from y to z_m. If $y = -fx + (1+f)z_m$ then $dy = (1+f)dz_m$. So, summing over all possible values of x by integrating, we get the probability of a female having phenotype x encountering a male with corresponding phenotype y as in (23). Then, multiplying by the number of females and the average number of male offspring per female, we get an equation representing the number of male offspring

with phenotype z_m

$$\tilde{m}_t(z_m)dz_m = \frac{r}{2}(1+f)\int n(x)\frac{m\left(-fx+(1+f)z_m\right)}{M}\,dx\,dz_m,$$

so the density of male offspring is given by

$$\tilde{m}_t(z_m) = \frac{r}{2}(1+f)\int n(x)\frac{m\left(-fx+(1+f)z_m\right)}{M}\,dx. \tag{25}$$

This means that while mate choice is random in terms of phenotypes, there is some density dependence involved and rare phenotypes are penalized. For example, a low number of males with a given phenotype means that fewer females are likely to mate with an individual of that phenotype, resulting in fewer offspring with that phenotype combination. On the other hand, males with a common phenotype are more abundant so more females will mate and produce offspring with individuals having that trait value. This means more common male phenotypes are more likely to affect the offspring phenotype distribution in the next generation.

To produce a female offspring with phenotype z_f we look at the probability that a female with phenotype x encounters a male with a corresponding phenotype y such that (24) is satisfied. Using a similar change of variables, we can rewrite $n(x)$ in terms of the female offspring phenotype z_f, where $x = -fy + (1+f)z_f$ so $dx = (1+f)dz_f$. As above, we integrate the product of these distributions over all possible y to end up with the density of female offspring,

$$\tilde{n}_t(z_f) = \frac{r}{2}(1+f)\int n\left(-fy+(1+f)z_f\right)\frac{m(y)}{M}\,dy, \tag{26}$$

where, again, the additional $(1+f)$ term is a normalizing constant that is the result of changing variables from the parent phenotype to the offspring phenotype.

Numerically, implementing (25) and (26) was a bit of a challenge. The basic model was implemented by specifying a vector of valid phenotypes from -1 to 1 with some predetermined stepsize, Δz. The density of these phenotypes was specified as a vector with the same stepsize. To keep the spacing of the offspring phenotypes the same, we needed to first

determine the phenotype of the parents that would result in offspring with phenotypes from -1 to 1 with spacing Δz, then we needed to determine the abundance of each parent phenotype in the population. For simplicity, we assumed that for the male offspring, the phenotype of the mother, x, lay along the predetermined spacing and had density $n_t(x)$. For the female offspring, the phenotype of the father, y, lay along that spacing and had density $m_t(y)$. Then, using (23) and (24), we computed the phenotypes of males and females respectively, that were necessary to produce offspring with phenotypes z_m and z_f for each value between -1 and 1.

Because equations (23) and (24) take the sum of two vectors with stepsize Δz that are multiplied by f and $1 + f$, there is no guarantee that the resulting vectors, $\hat{y}$ and $\hat{x}$, will also have that spacing. This means that we don't have predefined density values for those intermediate phenotypes. To determine the abundance of these intermediate phenotypes $\hat{y}$ and $\hat{x}$, we used one-dimensional linear interpolation to calculate the intermediate values for our distribution. Another result of equations (23) and (24) is that some of the phenotypes $\hat{y}$ and $\hat{x}$ that could mathematically produce offspring between -1 and 1 are not allowable in the context of the model because they lay outside the allowable range of phenotypes. To correct for this, we specified dummy phenotype variables that extended from -2 to 2 instead of just -1 to 1 and then redefined the distributions $n_t(z)$ and $m_t(z)$ so that there was zero probability of an individual existing outside the phenotype range -1 to 1, and the probability for a phenotype within this range was the same as originally specified. This allowed us to interpolate to find the density of all of the parent phenotypes. Finally, as another precaution against allowing individuals outside the phenotype range of -1 to 1, we again set the interpolated densities to zero for out of bounds phenotypes. Then, to compute equations (25) and (26), we computed the inner product of the female density and the male distribution, then used trapezoidal quadrature to integrate and obtain the offspring densities, $\tilde{m}_t(z)$ and $\tilde{n}_t(z)$.

Now that we have obtained the male and female offspring densities, this model proceeds

as in the clonal case. Mutation is represented via a convolution of the offspring densities with the Gaussian mutation distribution defined in (2). Then, after determining the post mutation densities, $\tilde{\tilde{m}}_t(z)$ and $\tilde{\tilde{n}}_t(z)$, we use (5) to compute the probability of each offspring surviving to adulthood, subject to the same ecological dynamics as in the clonal case. This results in the densities $m_{t+1}(z)$ and $n_{t+1}(z)$, which describe the population of males and females in generation $t+1$ who have survived to adulthood but who have not yet mated. From here, the process outlined in Figure 1 begins again.

5 Effective Particle Theory

In the clonal and sexual reproduction models developed in Sections 2 and 4, we began with the assumption that at the beginning of generation t, we had a population of M_t males and N_t females whose phenotypes could be described by some distribution. This general assumption allowed analysis of population dynamics and parameter effects using numerical methods, however the generality made the process intractable analytically. If however, we restrict the form of the population distributions to be only Gaussian, and allow the traits to range from $-\infty$ to ∞ for the purposes of integration, then we are able to use a technique called *Effective Particle Theory* (EPT) or the "Method of Moments" to approximate these populations in subsequent generations using Gaussian distributions. This method consisted of approximating the first three moments (mass or population, mean and variance) of the male and female phenotype distributions using three integrals described below. The intent of this method is to allow the use of analytic techniques which could feasibly allow for stability analyses of a nonmonomorphic population. This more general analysis could yield more realistic results than the stability analysis included in Section 3. Another goal is to reduce the dimension of the system from infinite-dimensional distributions in each generation to just six parameters, the populations of males and females, the means of each distribution, and their variances.

To begin, we assume that the males and females in generation t who have survived

to adulthood but who have not yet mated can be described by the following Gaussian distributions

$$m_t(z) = \frac{M_t}{\sqrt{2\pi\sigma_m^2}} \exp\left\{-\frac{(z-\mu_m)^2}{2\sigma_m^2}\right\} \tag{27}$$

and

$$n_t(z) = \frac{N_t}{\sqrt{2\pi\sigma_f^2}} \exp\left\{-\frac{(z-\mu_f)^2}{2\sigma_f^2}\right\}, \tag{28}$$

where M_t is the number of males, N_t is the number of females, μ_m and μ_f are the male and female means, and σ_m^2 and σ_f^2 are the male and female variances respectively.

As we go through the life cycle, the populations mate using either clonal or sexual reproduction, mutate as described in Section 2, and then a portion of these mutated offspring survive to adulthood based on (5). Because we have restricted the original population distributions to Gaussian forms, we are able to analytically compute each step of this process using standard Gaussian integration techniques. Evaluating the steps in the life cycle, we begin with reproduction. Clonal reproduction ($f = 0$) causes no change in the phenotypes and only requires the number of individuals to be updated, so that

$$\tilde{N}_t = \tilde{M}_t = \frac{rN_t}{2}$$

is true. In the sexual reproduction case, mating must be calculated as the integral of the product of Gaussian distributions, including a change of variables, as outlined in (25) and (26). Similarly, mutation can be evaluated analytically as the integral of a product of Gaussians, and as such, the mean and variance parameters remain meaningful and easily understood from the distribution.

Following mutation however, we wish to determine the probability that an offspring will survive competition and the effects of carrying capacity to adulthood so that they can also produce offspring. This is done by multiplying the offspring distributions after mutation, $\tilde{\tilde{m}}_t$ and $\tilde{\tilde{n}}_t$, with the probability of survival equation, (5), which results in a sum of Gaussian distributions which is itself non-Gaussian. Iterating through further generations would be possible numerically, though the resulting distributions would quickly become difficult to understand without graphing and additional computations.

This is where EPT is useful. We can determine the population numbers, means and variances in the next generation by calculating the first three moments, (Note, in equation (30), the μ_{t+1} refers to the mean phenotype of the population, not the mutation rate.)

$$N_{t+1} = \int \tilde{\tilde{n}}_t(z)P(z)dz, \tag{29}$$

$$\mu_{t+1} = \int z\tilde{\tilde{n}}_t(z)P(z)\frac{dz}{N_{t+1}}, \quad \text{and} \tag{30}$$

$$\sigma^2_{t+1} = \int (z-\mu_{t+1})^2\tilde{\tilde{n}}_t(z)P(z)\frac{dz}{N_{t+1}}. \tag{31}$$

Due to the form of the probability of survival equation (5), each of the above integrals must be calculated as the sum of three integrals by breaking $P(z)$ into three parts,

$$P(z) = 1 - \frac{r-1}{r}\frac{n_{eff,m}(z)}{K(z)} - \frac{r-1}{r}\frac{n_{eff,f}(z)}{K(z)},$$

where $n_{eff,m}$ is the effective male population and $n_{eff,f}$ is the effective female population. This is slightly more complex computationally, but results in integrals that can be computed analytically. This was a main reason that the logistic probability of survival equation was used instead of the Beverton-Holt equation used in the Bolnick and Doebeli (2003) model (see Section 2.3).

Using the moments computed in equations (29)-(31), the approximate distributions for males and females at the beginning of generation $t+1$ are

$$m_{t+1}(z) \doteq \frac{M_{t+1}}{\sqrt{2\pi\sigma^2_{m,t+1}}} \exp\left[-\frac{(z-\mu_{m,t+1})^2}{2\sigma^2_{m,t+1}}\right] \tag{32}$$

and

$$n_{t+1}(z) \doteq \frac{N_{t+1}}{\sqrt{2\pi\sigma^2_{f,t+1}}} \exp\left[-\frac{(z-\mu_{f,t+1})^2}{2\sigma^2_{f,t+1}}\right]. \tag{33}$$

Because they are Gaussian, we can repeat this process indefinitely to determine the population dynamics in future generations. Figure 6 demonstrates how a non-Gaussian distribution can be approximated in this way with a Gaussian distribution.

Now, implementing the process outlined above to approximate the male and female densities with Gaussian distributions using both clonal and sexual reproduction, we will discuss

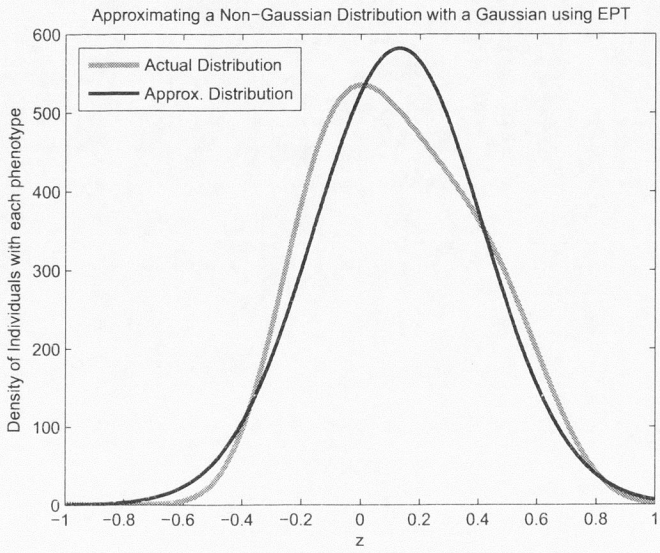

Figure 6: A non-Gaussian distribution (gray) is approximated by a Gaussian distribution (black) having the same population, mean and variance using EPT.

how the models performed and compare the full density-based phenotype model to the analytic EPT approximation.

6 Results

6.1 Numerical Results for Clonal Model

Using MATLAB to implement the clonal model described in Section 2, where we assume no shared material between genders ($f = 0$), we readily observed sexual dimorphism or evolutionary branching depending on the initial conditions of the population and the parameters associated with competition and carrying capacity.

When the variance of carrying capacity was less than the variance of competition, $\sigma_k < \sigma_c$, we observed that the effect of carrying capacity, the decrease in resources away from the optimum phenotype z_0, outweighed the effect of competition and the male and female mean phenotypes evolved toward the optimum as predicted by the stability analysis in Section 3. In this case, because there was still some competition, the means of the males and females did

not completely converge, but remained on either side of the optimum phenotype with some overlap between the genders. This can be seen in Figure 7 where the density of individuals is plotted over time. In these figures, black indicates phenotype regions with high population densities whereas shades of gray to white indicate phenotype regions with lower population densities.

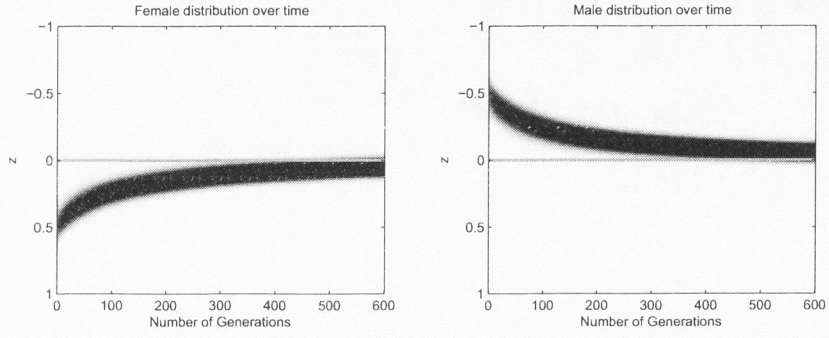

Figure 7: Density plot of male (right) and female (left) populations over time where $\sigma_k = 0.5 < \sigma_c = 1$. Dark gray to black regions indicate a high number of individuals in a particular phenotype range whereas light gray to white regions indicate few or no individuals in those phenotype ranges. The horizontal line in each figure represents the optimum phenotype $z_0 = 0$ where resources are most abundant. In these figures, we see that males and females evolve toward the optimum phenotype as predicted by the stability analysis in Section 3. Starting parameters were $r = 3, N = M = 200, K_0 = 2000, \mu_m = -0.5, \mu_f = 0.5, \sigma_m = \sigma_f = 0.1$, and $\sigma_\omega = 0.002$.

On the other hand, if we restrict ourselves to the case where $\sigma_k > \sigma_c$, then we observe more interesting things. If we begin with males and females having initial means near, but on opposite sides of, the optimum phenotype and small variance in the population, then we see that the effect of competition outweighs the benefits of being near the optimum phenotype. This causes the males and females to quickly evolve away from each other to more stable phenotypes where they have reached some balance between optimal resources and reduced competition. Over time, they will actually evolve toward the stable phenotypes described by (20) that we determined in Section 3. This situation, where male and female phenotypes become more different over time could be classified as sexual dimorphism. By changing the initial conditions slightly however, such as by moving the initial means closer to each other

or increasing the variance of the population, we can observe a different type of behavior where the phenotype range of one or both genders becomes more broad and bimodal and ultimately splits into two distinct ranges. This indicates evolutionary branching and allows the offspring, over time, to move toward phenotype ranges with less severe competition. This is shown in Figure 9. In cases of evolutionary branching where one or both gender splits into distinct phenotype ranges with few or no individuals between them, than this could be interpreted as a possible speciation event.

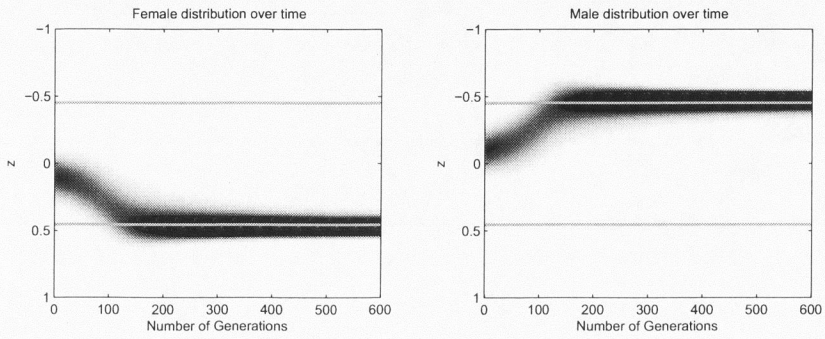

Figure 8: Density plot of male (right) and female (left) populations over time where males and females begin with $\mu_m = -0.1$, $\mu_f = 0.1, \sigma_m = \sigma_f = 0.1$, $r = 3$, $K_0 = 2000$, $z_0 = 0, \sigma_k = 1, \sigma_c = 0.85$, and $\sigma_\omega = 0.002$. Again, dark gray to black regions indicate a high number of individuals in a particular phenotype range whereas light gray to white regions indicate few or no individuals in those phenotype ranges. In this figure, we see a dimorphism occurs as the males and females evolve away from each other toward the horizontal lines representing the antisymmetric fixed points computed in Section 3.

So we see, using the density-based clonal model developed in Section 2, we are able to reproduce both sexual dimorphism (Figure 8) and evolutionary branching (Figure 9). We note, however, that the clonal model does not behave exactly as predicted by the stability analysis. Often, it exhibits more complicated behavior including additional evolutionary branching near the asymmetric fixed points. An example can be seen in Figure 10. This alone indicates the need for a density-based model because stability analyses assuming monomorphic populations simply cannot capture all of the complexity involved when trying to represent the dynamics of the entire population.

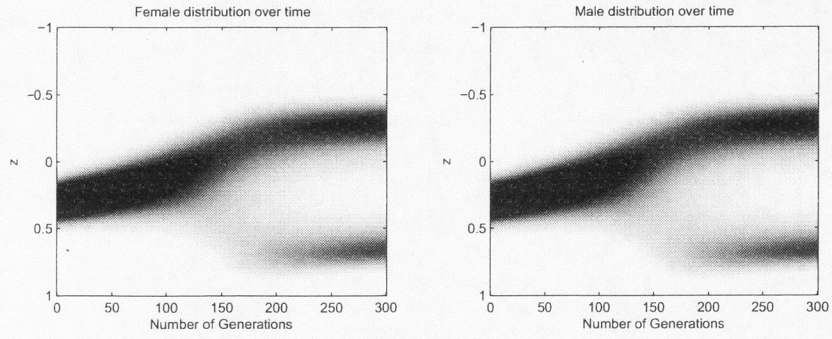

Figure 9: Density of male (right) and female (left) populations over time where males and females began with $\mu_m = \mu_f = 0.3$ and $\sigma_m = \sigma_f = 0.1$, $r = 3, K_0 = 2000, z_0 = 0, \sigma_k = 1, \sigma_c = 0.85$ and $\sigma_\omega = 0.005$. Again, dark gray to black regions indicate a high number of individuals in a particular phenotype range whereas light gray to white regions indicate few or no individuals. In this figure, both sexes tend toward the optimum phenotype ($z_0 = 0$) until they experience a disruptive event near generation 120. This causes each sex to exhibit evolutionary branching by splitting into a bimodal distribution (indicating possible speciation).

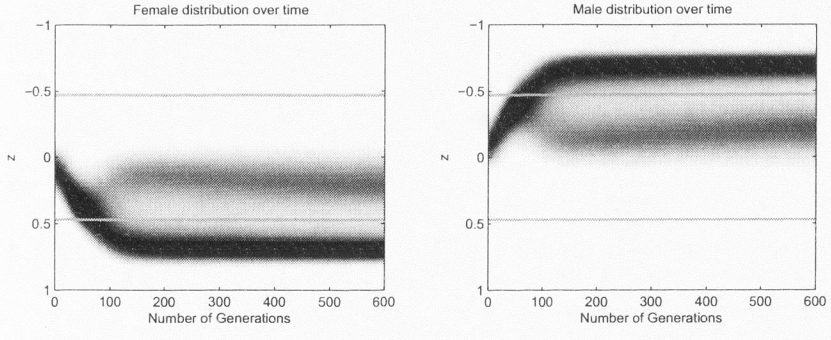

Figure 10: Density of male (right) and female (left) populations over time where males and females began with $\mu_m = -0.1, \mu_f = 0.1, \sigma_m = \sigma_f = 0.1$, $r = 3, K_0 = 2000, z_0 = 0, \sigma_k = 1, \sigma_c = 0.45$ and $\sigma_\omega = 0.002$. Similar to Figure 8, sexual dimorphism again results from high competition, however as the genders evolve toward the horizontal lines representing the asymmetric fixed points, a second disruptive event occurs and the distribution of each sex splits apart. This is an example of sexual dimorphism followed by evolutionary branching.

6.2 Analytic Approximation of EPT and Comparison with Numerical Results for the Clonal Model

Using *Mathematica* to implement EPT for the clonal model, we arrived at six equations representing the population, mean and variance of the males and females in generation $t+1$

based on the population, mean and variance in generation t as well as the parameters describing carrying capacity and competition. Unfortunately, these equations were excessively complicated. The simplest is the equation for the mass (number) of the populations. The expression for the female population is shown below,

$$N_{t+1} = \frac{rN_t}{2} - \frac{e^{-\frac{\mu_f^2\left(\sigma_c^2+2\left(\sigma_f^2+\sigma_\mu^2\right)\right)}{2\left(\sigma_f^4+\sigma_\mu^2(-2+\sigma_\mu^2)+2\sigma_f^2(-1+\sigma_\mu^2)+\sigma_c^2\left(-1+\sigma_f^2+\sigma_\mu^2\right)\right)}}(-1+r)rN_t^2\sigma_c}{4K_0\sqrt{-\sigma_f^4-\sigma_\mu^2\left(-2+\sigma_\mu^2\right)-2\sigma_f^2\left(-1+\sigma_\mu^2\right)-\sigma_c^2\left(-1+\sigma_f^2+\sigma_\mu^2\right)}} \tag{34}$$

$$-\frac{e^{-\frac{2\mu_f\mu_m+\mu_m^2\left(-1+\sigma_f^2+\sigma_\mu^2\right)+\mu_f^2\left(-1+\sigma_c^2+\sigma_m^2+\sigma_\mu^2\right)}{2\left(-\sigma_m^2-2\sigma_\mu^2+\sigma_m^2\sigma_\mu^2+\sigma_\mu^4+\sigma_c^2\left(-1+\sigma_f^2+\sigma_\mu^2\right)+\sigma_f^2(-1+\sigma_m^2+\sigma_\mu^2)\right)}}(-1+r)rN_t^2\sigma_c}{4K_0\sqrt{\sigma_m^2+2\sigma_\mu^2-\sigma_m^2\sigma_\mu^2-\sigma_\mu^4-\sigma_c^2\left(-1+\sigma_f^2+\sigma_\mu^2\right)-\sigma_f^2\left(-1+\sigma_m^2+\sigma_\mu^2\right)}}.$$

Due to the nature of the moment equations, the equations for the mean and variance became far more complicated still. The equation for the female mean using clonal reproduction was three times as long as the above equation, and the equation representing the female variance cannot be fit onto a single page. So, while in theory, a stability analysis could be done on these six equations, it was too comprehensive a task to attempt here. Instead, we utilized these equations to iteratively compute the population, mean and variance for future generations and then compared those values to the numerical results for the full density-based model calculated using MATLAB.

Utilizing numerous initial parameter combinations, we found that for combinations resulting in just dimorphism, the means computed numerically are approximately the same as the means computed using EPT (see Figure 11, graphs A and B). We do note, however, that the analytic mean can take slightly longer to converge to the stable phenotype, and, due to differences between the numerical and analytic computations, the rate of mutation, σ_ω, needed to be increased for the EPT approximation computed by *Mathematica*.

For parameter combinations resulting in evolutionary branching, however, EPT is not predictive. This is because EPT is compiling all the data from the distribution into a Gaussian distribution, approximating the true distribution with a unimodal distribution, and hence is not able to represent bimodal distributions. Similarly, in cases of extreme competi-

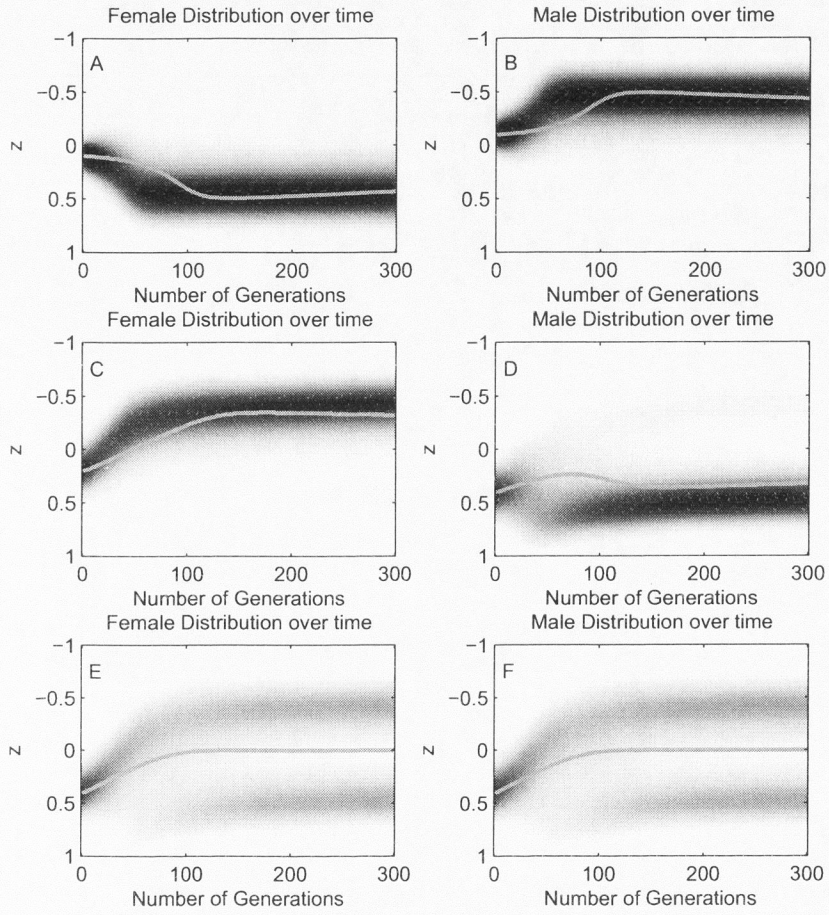

Figure 11: Comparison of male (right) and female (left) distributions over time using the full numerical Clonal model (density plot) and computed EPT means (overlaid lines). In all figures, $r = 3, K_0 = 2000, z_0 = 0, \sigma_k = 1$ and $\sigma_\omega = 0.002$. In figures A and B, males and females begin with $\mu_m = -0.1$, $\mu_f = 0.1, \sigma_m = \sigma_f = 0.1$, and $\sigma_c = 0.7$; in this example, neither the male nor female distribution splits apart, and the line representing the EPT mean is a good approximation. In figures C and D, we begin with $\mu_m = 0.4$, $\mu_f = 0.2, \sigma_m = \sigma_f = 0.15$, and $\sigma_c = 0.8$. Here, the male distribution experiences some evolutionary branching (splitting, indicating possible speciation) and the line representing the EPT mean for males is less accurate. In figures E and F, we begin with $\mu_m = 0.4$, $\mu_f = 0.4, \sigma_m = \sigma_f = 0.15$, and $\sigma_c = 0.8$. In this case, both genders exhibit evolutionary branching (possible speciation) and the EPT mean lies in the middle of the entire distribution and does not provide an accurate approximation of the truth.

tion, when σ_c is very small, the numeric calculations of the distributions can result in even more modes (see Figure 10), and EPT again cannot represent this complexity. Comparisons of the numeric and analytic predictions for male and female phenotypes are shown in Figure 11.

Comparisons of the male and female populations and variances showed that, for parameter combinations not leading to evolutionary branching, the values estimated by EPT agreed well. EPT often predicted wider variances than the full density-based model, with values converging to approximately 0.5 versus numerical variances that converged more often to values between 0.1 and 0.2. This however, could be explained by the fact that when implementing EPT, parameter values were allowed to range from $-\infty$ to ∞ versus -1 to 1 in the full density-based model. For parameter combinations leading to more complicated branching behavior, the EPT values became less accurate. This loss of accuracy was likely due to the reliance on each parameter in computing the population, mean and variance in the next generation. As the mean parameter estimation became less accurate, this loss of accuracy propagated through, leading to less accurate values for the population and variance.

6.3 Effect of varying σ_ω

Analyzing the effect of mutation, we saw in Figure 2 that mutation has the effect of adding in variance to the population. The larger the variance of $\Omega(z, z')$, the mutation distribution, the more variance gets injected into the population at each time step and the more spread out the phenotypes become. In some cases, this allows for a small group of individuals to escape intense competition or low carrying capacity and reach an area of the fitness landscape where they and their offspring will have a better chance of surviving. This is what allows for the dimorphism shown in Figure 8 and the evolutionary branching shown in Figure 9.

We utilize a small mutation variance for most iterations because mutation probabilities are generally small. However, due to limitations in the numerical computation of the models which will be discussed in Section 7.1, we are limited as to what values of σ_ω we are able to use. To see how different values of μ or σ_ω affect a population over time, we fixed all

parameters except σ_ω and plotted a female distribution over time with various (increasing) values of σ_ω. This is shown in Figure 12. What we notice in these graphs, is that without mutation, sexual reproduction causes the male and female distributions to evolve toward each other, when some mutation is added, males and females are able to respond to the stronger competition by branching into two phenotype ranges less affected by competition. Larger values of σ_ω have the effect of spreading the range of phenotypes to the point where intraspecific competition has a lesser effect on individuals, and reduces their need to spread to distinct phenotype ranges. As σ_ω gets too large however, we also begin to experience wraparound error due to the numerical implementation of mutation. Similarly, if σ_ω is too small, we experience other numerical errors. These errors will be discussed further in Section 7.1.

In the absence of mutation, depending on starting parameters, we may still see individuals escape extremes of the fitness landscape provided the variance of the initial male and female distributions are wide enough that initially, some males and females were less affected by competition and low carrying capacity. If the initial variance of the distributions is not wide enough however, then in the absence of mutation, we will generally not see a change in the mean phenotype over time.

6.4 Numerical Results for Sexual Reproduction Model

Implementing sexual reproduction with random mating (where we assumed there was shared material between the genders), we observed that as the ratio of shared material between the genders, f, was increased, the chances for sexual dimorphism and speciation were strictly limited. As in the clonal case, these dynamics were complicated, depending on initial population parameters for both genders, the influence of carrying capacity and competition, mutation and f, the ratio of shared genetic material between genders. Due to this large number of parameters influencing competition, it was difficult to determine how they interacted.

To explore these dynamics, we began with the assumption that $\sigma_c < \sigma_k$ and began with no mutation and an extremely low ratio of shared material, $f = 0.02$. Using a male

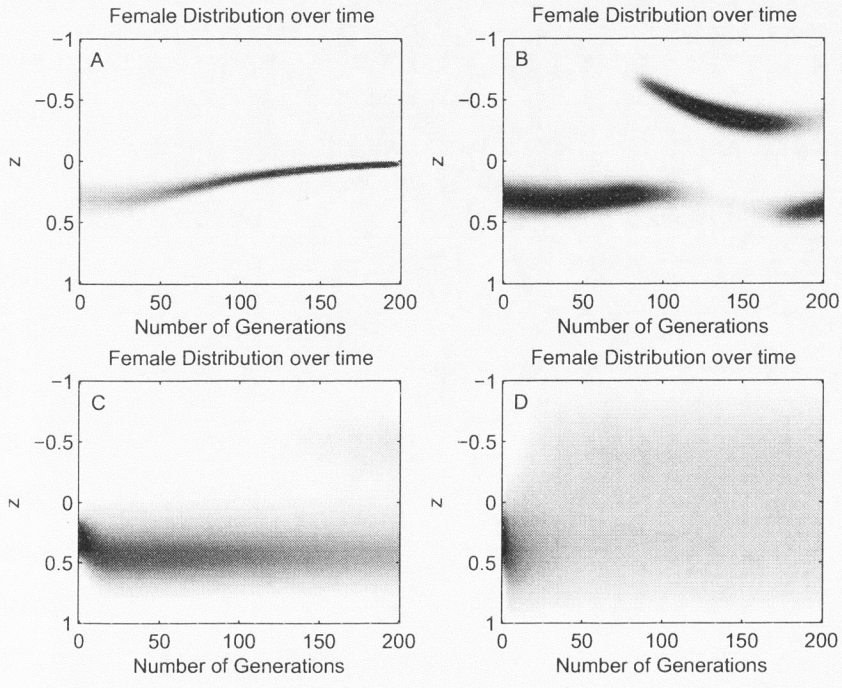

Figure 12: Density plot of female distribution over time (using sexual reproduction) with various values of the mutation parameter σ_μ. These graphs show how different mutation parameter values can affect the numerical computation of the density-based model. In graphs A-D, starting values were $r = 4, \mu_m = 0.3, \mu_f = 0.3, \sigma_m = 0.1, \sigma_f = 0.1, K_0 = 2000, \sigma_k = 1, \sigma_c = 0.65$. In graph A, there is no mutation, in graphs B-D, the mutation parameter, σ_ω, is increased to 0.01, 0.05, and 0.12. Graphs C and D show how larger values of σ_ω allow the distribution to spread toward the boundary. This is most extreme in graph D where serious wrap around error is shown. This type of error can lead to unrealistic results.

population with initial mean $\mu_m = -0.3$ and females with initial mean $\mu_f = 0.3$ we saw that mating and shared material had the effect of forcing the males and females over time to the same mean with almost no variance in the population. Adding in a very small amount of mutation $\sigma_\omega = 0.01$, we saw the male and female mean merge toward zero again, however as the two distributions met near generation 100, they temporarily displaced each other to a degree causing a slight speciation or dimorphic event, but ultimately the disruptive event was not enough to stabilize as a dimorphism or speciation and the populations re-merged. Increasing the mutation again, to $\sigma_\omega = 0.015$ we see a similar occurrence. The males and

females merged toward zero, experienced a disruptive event near generation 100 and then split into two modes. This time however, the mutation rate is great enough to sustain this speciation for a period of time, resulting in a bimodal population of males and females with modes on either side of the optimum phenotype $z_0 = 0$. This behavior is shown in Figure 13. As σ_ω, the variance of the mutation distribution, is increased further, this trend continues: the populations of males and females become bimodal until the mutation distribution gets too wide, flattening the competition curve, allowing the males and females to spread more uniformly as seen in Figure 12. This occurs when mutation is at an extremely high and unrealistic rate, so we will return to the case where we allow a small amount of mutation σ_ω and instead vary f.

In Figure 14 we fix the rate of mutation at a low level, $\sigma_\omega = 0.02$, and explore the effect of the parameter f. From the clonal model, we know that when $f = 0$, depending on initial parameter combinations, both sexual dimorphism and evolutionary branching can occur. Using a low ratio of shared material, $f = 0.01$, we also observe evolutionary branching. By increasing the ratio to 0.15, we observe that the evolutionary branching occurs sooner, around generation 100 instead of 150. After increasing the ratio of shared material again to $f = 0.03$ we no longer see speciation, and instead see the male and female distributions overlap. From here, further increases of f will have the effect of causing the genders to share more and more material, causing their offspring to become more and more similar. These distributions will simply converge to the optimum phenotype and do not display any type of speciation or dimorphism unless mutation is again increased to an unrealistic level.

6.5 Sexual Reproduction - Comparing Numerical Results to Effective Particle Theory

As in the clonal case, the EPT is reducing the dimension of the data to just six parameters, the number of males and females, mean of the male and female phenotype distribution, and the male and female variance at each time step. As such, it is unable to predict situations where the males or females become dimorphic as was demonstrated for extremely low values

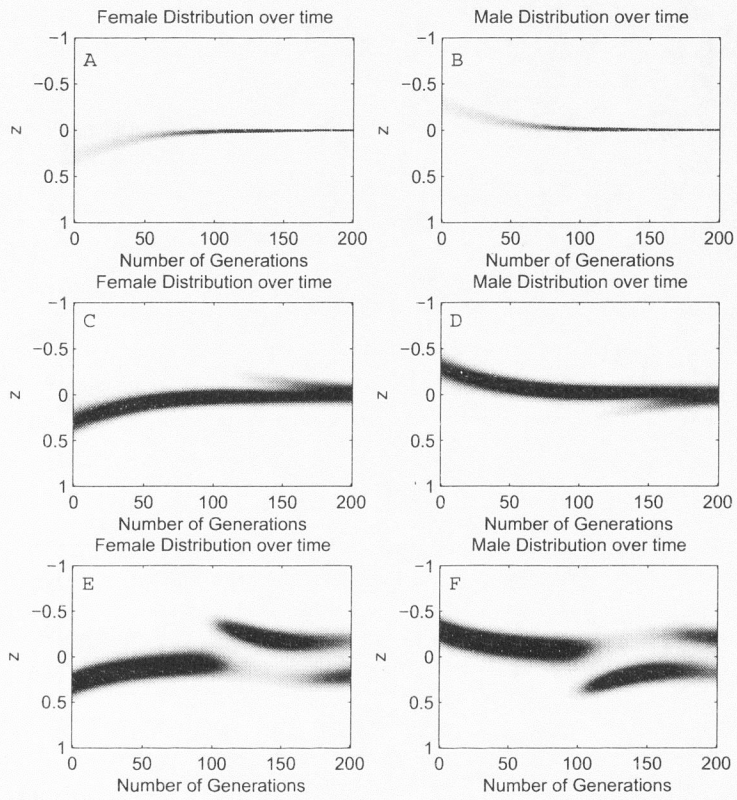

Figure 13: Density of male (right) and female (left) distributions over time using sexual reproduction with various mutation parameters. In images A and B, there is no mutation and the male and female populations become nearly identical to each other over time, in images C and D, $\sigma_\omega = 0.01$ allowing the male and female distributions to be more broad and temporarily bimodal, and in images E and F, the mutation parameter, $\sigma_\omega = 0.015$, is large enough to allow the male and female distributions to become bimodal and exhibit evolutionary branching. In all figures, $f = 0.02, r = 4, z_0 = 0, K_0 = 1000, \sigma_k = 1, \sigma_c = 0.7, N = 200, M = 200, \mu_n = 0.3, \mu_m = -0.3$, and $\sigma_f = \sigma_m = 0.1$.

of f in Section 6.4. For larger values of f however, when males and females simply converge to the optimum phenotype, EPT was able to accurately predict this behavior. This can be seen in Figure 15.

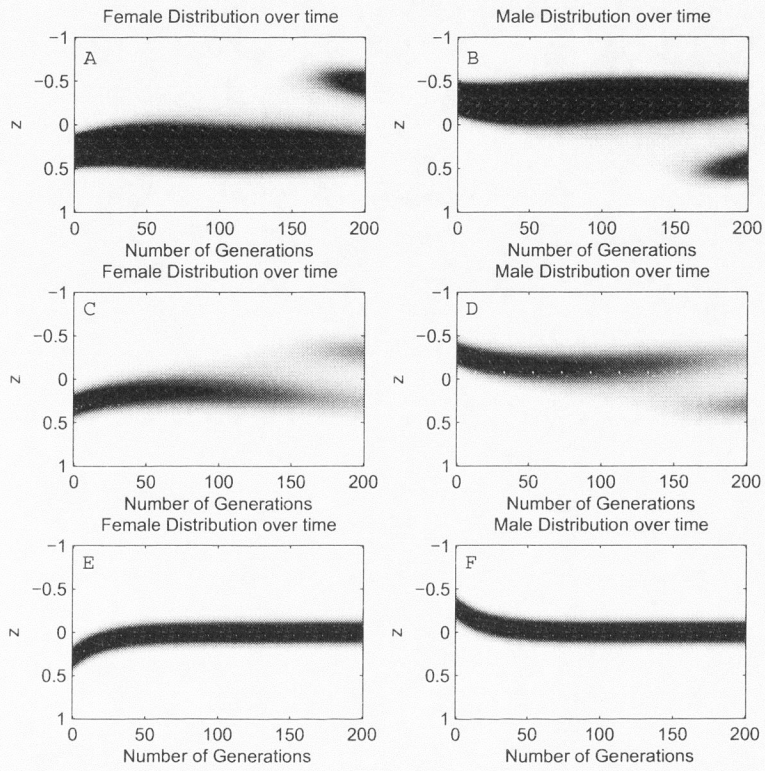

Figure 14: Exploring the effect of the ratio of shared material by comparing the density of male (right) and female (left) populations over time with various values of f. Darker regions indicate a higher number of individuals in a phenotype range, light gray to white regions indicate few or no individuals. Initial parameters for all figures starting values $r = 3, z_0 = 0, K_0 = 1000, \sigma_k = 1, \sigma_c = 0.7, \sigma_\omega = 0.02, N = 200, M = 200, \mu_n = 0.3, \mu_m = -0.3$, and $\sigma_m = \sigma_f = 0.1$. In figures A and B, $f = 0.01$ meaning there is very little shared material and the populations stay quite different and eventually exhibit evolutionary branching. In figures C and D, $f = 0.015$, so there is more shared material causing the populations to tend toward each other yet still branch, and in figures E and F, $f = 0.03$. In these figures, there is too much shared material for the populations to remain in distinct phenotype ranges.

7 Discussion

In comparing the density-based models developed in Sections 2 and 4, we found that the models behaved in a similar manner to the individual-based models developed by Bolnick and Doebeli (2003). In the clonal model, slight differences existed regarding fixed points and stability due to the use of the logistic probability of survival equation, but the density-based

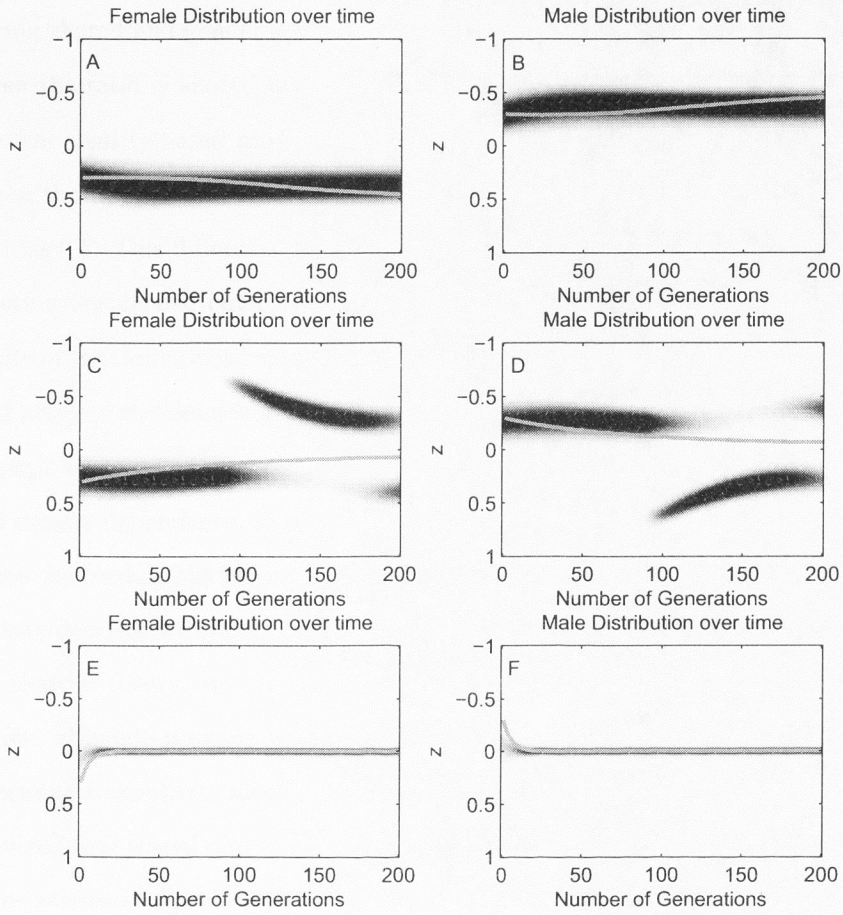

Figure 15: Comparison of male (right) and female (left) distributions over time using the full numerical sexual reproduction model (density plot) and computed EPT means (overlaid lines). In figures A and B, $f = 0.005$ and neither sex exhibits evolutionary branching. We see that in this case, the overlaid EPT mean approximation is similar to the mean calculated using the full numerical model. In figures C and D, $f = 0.01$ and the sexes both exhibit evolutionary branching. Because both sexes are now represented with bimodal distributions, the single EPT mean is not able to provide a good approximation. Finally, in figures E and F, $f = 0.1$, and the high ratio of shared material causes the sexes to become nearly identical. Since there is no evolutionary branching, the EPT mean is a good approximation. Other parameter values in these figures are $r = 4, z_0 = 0, K_0 = 1000, \sigma_k = 1, \sigma_c = 0.65, \sigma_\omega = 0.01, N = 200, M = 200, \mu_f = 0.3, \mu_m = -0.3, \sigma_n = 0.1, \sigma_m = 0.1$.

model was able to predict both sexual dimorphism and speciation. For the model including sexual reproduction, the main difference was due to the way mating was implemented

mathematically. We utilized f, the ratio of shared material between genders rather than considering shared and gender specific alleles. As a result, sexual dimorphism was observed only when the ratio of shared material was extremely low. For example, to witness dimorphism in the density-based model, f needed to be on the order of one half of one percent whereas in the Bolnick and Doebeli (2003) model, they witnessed dimorphism for values of f as high as 0.5. This difference however, may be due to the different ways in which sexual reproduction was carried out in each model. The ratios of shared material do not have the same units in the individual-based model as in the density-based model. A second difference we noted between the Bolnick and Doebeli (2003) model and the density-based model is that even though we specified mating as random from a phenotype selection standpoint, we still included density dependence, that is, the density of males with a particular phenotype determined how successful that phenotype would be in creating offspring in the next generation. This density dependence allowed for speciation to occur with certain parameter combinations showing that, in theory, assortative mating is not strictly necessary for speciation provided the specific, although possibly unlikely, parameter combinations exist.

Performing a sensitivity analysis on σ_ω in both the clonal and sexual reproduction model, we discovered that sexual dimorphism or speciation was most dependent on the initial conditions. The existence of particular outcomes depended on the variance of the carrying capacity σ_k^2 and the variance of the competition function σ_c^2. The relative sizes of these parameters decided whether the population was most likely to converge to the optimum phenotype - indicating that the carrying capacity of the environment had the greatest effect on them, or whether they were most affected by competition with each other - indicating that the populations would need to evolve away from each other to survive. Then, depending on the width of the initial male and female distributions and their initial means, the populations would either evolve toward or away from each other. If the initial distributions of males and females were wide enough, so that individuals existed in phenotype ranges that were not excessively hindered by competition, the populations were still able to spread apart and branch without

mutation. On the other hand, in the absence of mutation, if the initial distributions were not wide enough, they would remain in a phenotype range despite excessive competition. As the mutation rate increased, the distributions became more spread out, flattening the effect of competition and eliminating the necessity to spread to distinct phenotype ranges.

In the mating case, sexual dimorphism and speciation were much more constrained (and unlikely in general) and it turned out that mutation was necessary for these outcomes. As in Figure 13, in the complete absence of mutation, the phenotypes of the males and females simply mixed and evolved toward each other, and the variance of the population became very low. As some mutation was added in, we saw the males and females evolved toward each other, then began to experience more competition due to high population density near the optimum phenotype. This caused a temporary disruptive event in which the male and female populations became bimodal for several generations, then converged back to a single mode. Increasing σ_ω slightly, we saw the disruptive event became more permanent, resulting in a bimodal population of males and females indicating speciation had possibly occurred. Increasing the mutation rate still further, we saw the male and female distributions become so wide that competition had less effect, and the distributions simply became overlapping, wide and unimodal so that neither speciation nor dimorphism occurred.

Analyzing the effect of f in the mating case, we saw that for everything but very small values, there was simply too much sharing of material between genders, and the male and female offspring quickly evolved to the same phenotype. Increasing f slowly from zero however, we see that the males and females will be dimorphic initially. As f is increased, the dimorphism destabilizes to speciation until f is increased to the point where the genders become too well mixed, rendering both speciation and dimorphism impossible.

So we see, in the sexual reproduction model, the presence of mutation is necessary for dimorphism and/or evolutionary branching to occur. If however, the rate of mutation gets too high, then mutation flattens the male and female distributions which flattens the effect of competition, removing the possibility for dimorphism. A similar thing happens with the

parameter f. When $f = 0$, we simply have the clonal case where both dimorphism and speciation can occur. As f is slowly increased, we now have shared material between the genders. Sexual dimorphism remains possible only for extremely small values of f and as the ratio of shared material is increased, dimorphism is replaced with speciation until too much material is shared causing the males and females to simply merge and become identical.

7.1 Parameter Limitations Due to Numerical Implementation

When specifying initial parameter combinations with which to run the density-based model numerically, there were certain limitations we needed to respect. First, due to the self imposed phenotype range of -1 to 1, we needed to ensure that when comparing numeric and analytic EPT results, we began with a distribution as close to Gaussian as possible. This required the use of small variances and means near the center of the phenotype range to ensure that only small amounts of the distribution were being truncated at the phenotype boundary. Nonetheless, the distributions used for numeric iterations were not quite Gaussian and led to differences between computed results from the full numerical model and EPT.

Other restrictions on parameter combinations were due to the periodicity requirements of the FFT and discretization for numerical implementation in MATLAB. These limited the range of widths we were able to specify for functions being convolved. Two common issues that arose, were "wrap around" error and noise that can be introduced when the mutation distribution's width is too narrow relative to the stepsize or resolution specified for the phenotype range.

The FFT views the distributions specified as a single period of an infinite sequence. This means that values on the left hand side of a distribution meet the values on the right side forming a repeating sequence. When transformed to Fourier space, the distributions become continuous functions and are multiplied together. A problem occurs if this product no longer fits in the original period. If this occurs, then when the distribution is transformed back into phenotype space, portions of the distribution may spill from one boundary to the other, creating wrap around error or aliasing (Smith, 1997-1998; Oppenheim, Willsky, and

Nawab, 1997). In the density-based model, wrap around error occurs when the mutation distribution is too wide and also when the distributions of the males and females are too near the phenotype boundary either due to large variance parameters or means that are too close to phenotype extremes. This results in phenotypes literally wrapping around the phenotype boundary, migrating instantly from one extreme of the phenotype range to the other, i.e. -1 to 1 or 1 to -1. Wrap around error results in behavior that is not realistic from the standpoint of either the model or nature. We attempt to avoid wraparound behavior by limiting the size of the variance of both the male and female distributions and the mutation distribution.

The second type of error associated with numerical implementation is noise introduced when one of the functions being convolved is too narrow relative to the stepsize specified for the phenotype. For example, when working with a phenotype range of -1 to 1, we must specify the number of discrete steps we wish to use between these values. Since we are using the FFT to compute convolutions, it is most efficient to have 2^n steps for some integer n. This results in a stepsize of 2^{1-n}. If, for example, we use $n = 8$, then we have 256 steps between -1 and 1 and our stepsize is only 0.0078, meaning that if we wished to specify a mutation distribution with a standard deviation of 0.01, then we would have only a few data points describing the distribution which could cause serious errors when computing the FFT or any numerical algorithm. These errors can show up as oscillations or striations where the predicted density is, and also by noise at the edges of the phenotype range.

Using the Nyquist Sampling Theorem, which states that a function can be uniquely determined by its samples if the sampling frequency is greater than twice the maximum frequency of the function, we are able to avoid these errors (Oppenheim, Willsky, and Nawab, 1997). For example, to specify a mutation distribution with standard deviation σ_ω, we need the inequality $2\sigma_\omega > 4\Delta z$ to be satisfied, meaning that $\Delta z < \frac{1}{2}\sigma_\omega$. This is shown in Figure 16. Using the previous example, if we wish to specify a mutation distribution with standard deviation of 0.01, then we need to have a stepsize no greater than 0.005. This means that

the resolution above, 2^8, is not great enough because it only yields a stepsize of 0.0078 which is too large. So, to use the mutation distribution, we would need to increase the resolution to 2^9.

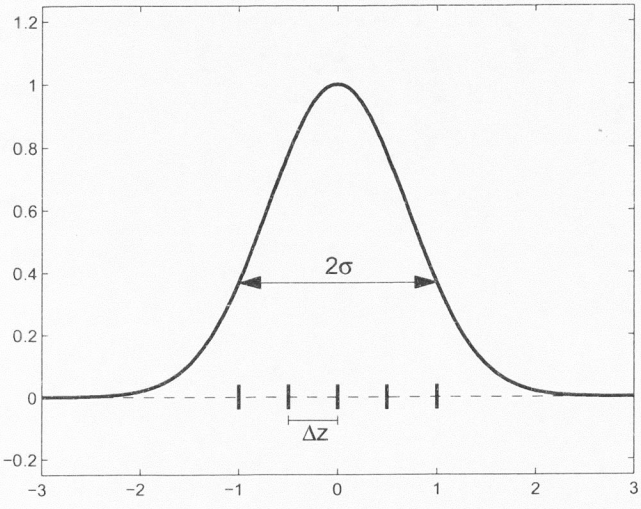

Figure 16: According to the Nyquist Sampling Theorem, we need to ensure that the sampling frequency (1/Deltaz) is at least two times the maximum frequency (1/sigma) of the function being represented.

Now that we understand the issues that can arise in implementing the models numerically, we will discuss a method for enhancing the usefulness and applicability of the analytic EPT method.

7.2 Extending Effective Particle Theory to Describe Bimodal Distributions

As discussed in Section 6.1, the EPT as described in Section 5 is not able to describe the behavior of population that becomes bimodal over time because it is reducing the dimension of the information describing each distribution at each time step to just three parameters: the population, mean and variance of a normal distribution. This obviously limits its usefulness, as one of the main behaviors of interest to ecologists is speciation. To determine whether

it was possible to adapt EPT to describe bimodal populations, we hypothesized that we might be able to describe the population using a sum of Gaussians. To determine these Gaussians, we would evaluate the population, mean and variance of the population on the phenotype interval $(-\infty, z_0)$ and (z_0, ∞) separately using equations (29)-(31) and then use these parameters to describe the Gaussians. Note that this split at phenotype zero was only imposed for the computation of the EPT parameters, and that the populations were free to roam the entire phenotype space over time. The idea was that with two Gaussians, hence two modes, we may be able to approximate the true behavior of the distributions more accurately in the case of evolutionary branching.

This would be a complex endeavor to evaluate analytically, so to test whether it would behave as expected, we chose implement the idea in MATLAB. Results for this method were not optimal and likely result from the computation of the variance. To improve the correlation between methods, we experimented with fixing the variance in the EPT computation or utilizing the variance calculated using a single EPT. The best representations seemed to come when the variance was fixed to one half the original variance for each Gaussian distribution. A comparison of the predicted future distributions comparing numerical implementation of the clonal model and numerical implementation of a double EPT for a population where both males and females exhibit evolutionary branching are shown in Figure 17.

In the case where neither gender exhibits evolutionary branching, we again had mixed results. For some initial parameter combinations, the double EPT is a decent representation as in Figure 18. In some other cases however, the double EPT predicted evolutionary branching of a small portion of the population, whereas direct numerical calculation did not. This indicates that EPT could possibly be used to model for dimorphism and/or evolutionary branching although more analysis into the best ways to approximate the variance parameter would be necessary.

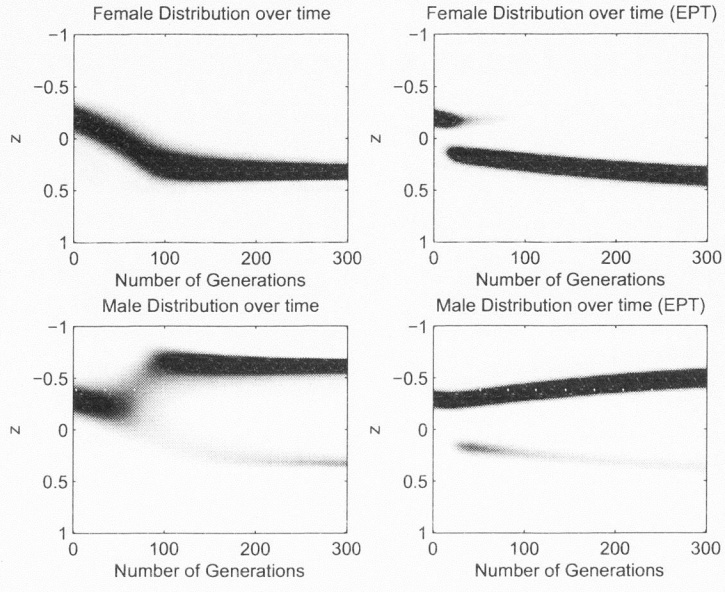

Figure 17: Comparison of the full numerical clonal model with double Effective Particle Theory. Females are shown on the top and males on the bottom with the starting values $\mu_m = -0.3, \mu_f = -0.2, \sigma_m = 0.1, \sigma_f = 0.1, K_0 = 2000, \sigma_k = 1, \sigma_c = 0.9, \sigma_\omega = 0.002$. Images on the left represent the numerical calculation and images on the right represent the results using a double EPT. For the female distribution, we see that there is no evolutionary branching, so EPT would already have provided a reasonable approximation. For the males, evolutionary branching does occur and over time, the double EPT approximation does mimic the full numeric projection indicating that this method may be capable of approximating more complicated behavior.

8 Conclusions

In this paper, we have adapted deterministic and stochastic models developed by Bolnick and Doebeli in "Sexual Dimorphism and Adaptive Speciation: Two Sides of the Same Ecological Coin" (Bolnick and Doebeli, 2003) into density-based mathematical models which describe the dynamics of a population of males and females using both clonal and sexual reproduction. These models characterize individuals according to trait value or phenotype and then determine the interactions of these phenotypes using a density-based approach. In each generation, individuals are subject to mating, the possibility for mutation, and ecological dynamics including effects of intraspecific competition and carrying capacity.

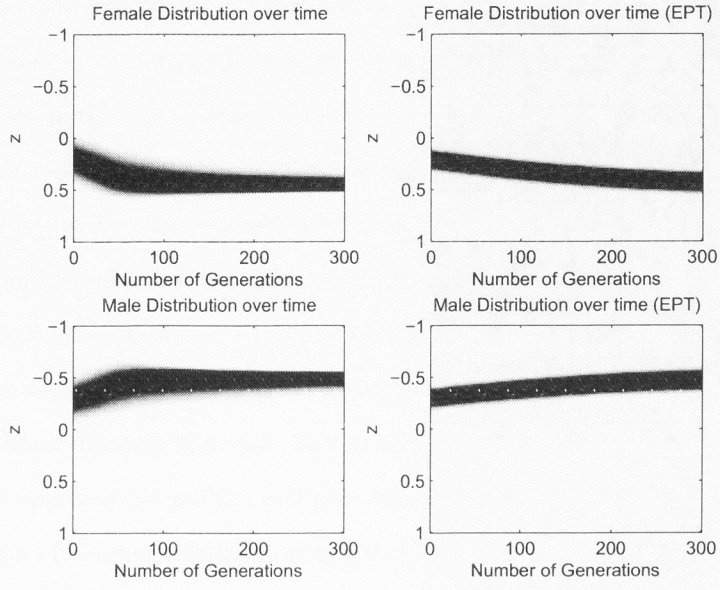

Figure 18: Comparison of the full numerical clonal model with Double Effective Particle Theory. Females are shown on the top and males on the bottom with the starting values $\mu_m = -0.3, \mu_f = 0.2, \sigma_m = 0.1, \sigma_f = 0.1, K_0 = 2000, \sigma_k = 1, \sigma_c = 0.9, \sigma_\omega = 0.002$. Images on the left represent the numerical calculation and images on the right represent the results using a double EPT. In these images, there is no evolutionary branching, and the double EPT projections are an accurate approximation.

Using these models, we have performed a stability analysis and explored the dynamics of rare mutants for the clonal density-based model when assuming a monomorphic population. We then demonstrated how numeric predictions computed by MATLAB using clonal reproduction match the expectations of the stability analysis in certain cases. We also discussed situations when the clonal model for a non-monomorphic population did not behave as predicted by the stability analysis. This demonstrated the need to perform a more complicated stability analysis of the density-based model because it seems that certain interactions are simply not apparent when assuming a monomorphic population.

Using MATLAB to implement the models, we demonstrated how sexual dimorphism and speciation could be readily observed when using clonal reproduction. We also showed how the sharing of material between genders, via sexual reproduction, tended to eliminate the chance

for dimorphism and speciation in cases where the ratio of shared material was too great. Nonetheless, the model did result in both sexual dimorphism and evolutionary branching with appropriate initial conditions.

To allow for more analytic study of sexual dimorphism and speciation, we demonstrated how EPT can be used to analytically predict the approximate population, mean and variance of males and females in subsequent generations by assuming that the populations at the beginning of each generation follow a Gaussian distribution. We also showed situations when the EPT estimate would not be a good approximation, such as when speciation occurred and the population became bimodal. To combat this, we discussed how the method could be adapted to approximate multimodal populations analytically by representing the populations as a sum of Gaussian distributions, although we acknowledge that this will require a substantial amount of computation.

Future work that could be done on this density-based approach would be to incorporate assortative mating as in Bolnick and Doebeli (2003), which would necessitate an additional convolution with a mate choice function. This would allow more complex behavior for more realistic proportions of shared material between genders. Requiring the mate choice function to be Gaussian, this could also be incorporated into the analytic model as well. Another area of the density-based model that could be improved is the interpretability with regards to genotypes. This model allows all ecological interactions and mating to occur at the phenotype level, however it is likely it could be adapted to allow more interpretability with regard to genotypes, although this would certainly add much more computational complexity.

The density-based model can be used to describe an entire population of males and females and their ecological interactions. While the majority of models developed to describe ecological interactions are individual-based and must be analyzed through the use of numerous lengthy stochastic simulations, a major benefit of the density-based approach is that it does not require these lengthy stochastic simulations. That is, the results are reproducible. A benefit of the density-based approach versus a deterministic approach for a monomorphic

population is that more complex results can be achieved when considering the density of phenotypes. The results are able to capture more complex behavior that is likely a better representation of what is occurring in nature.

References

Daniel I. Bolnick. Can intraspecific competition drive disruptive selection? an experimental test in natural populations of sticklebacks. *Evolution*, 2004.

Daniel I. Bolnick and Michael Doebeli. Sexual dimorphism and adaptive speciation: Two sides of the same ecological coin. *Evolution*, 57(11):2433–2449, November 2003.

Reinhard Burger, Kristan A. Schneider, and Martin Willensdorfer. The conditions for speciation through intraspecific competition. *Evolution*, 2006.

Michael Doebeli and Ulf Dieckmann. Evolutionary branching and sympatric speciation caused by different types of ecological interactions. *The American Naturalist*, 156:S77–S101, 2000.

Mark Kirkpatric and Virginie Ravigne. Speciation by natural and sexual selection: Models and experiments. *The American Naturalist*, 159:S1–S31, March 2002.

Mark Kirkpatrick and Scott L. Nuismer. Sexual selection can constrain sympatric speciation. *The Royal Society*, February 2004.

J. David Logan. *Applied Mathematics*. Wiley, third edition edition, 2006.

Alan V. Oppenheim, Alan S. Willsky, and S. Hamid Nawab. *Signals and Systems*. Prentice Hall, 1997.

John A. Rice. *Mathematical Statistics and Data Analysis*. Duxbury, 2007.

Montgomery Slatkin. Ecological causes of sexual dimorphism. *Evolution*, 1984.

Steven W. Smith. *The Scientist and Engineer's Guide to Digital Signal Processing*. 1997-1998.

G. Sander van Doorn, Ulf Dieckmann, and Franz Weissing. Sympatric speciation by sexual selection: A critical reevaluation. *The American Naturalist*, 163(5):709–725, May 2004.